THE NEURO-COGNITIVE PROFILE OF THEORY OF MIND IN HEALTHY AGEING

FOYZUL RAHMAN

Doctor of Philosophy

ASTON UNIVERSITY

MAY 2021

Aston University

Thesis Abstract

Theory of mind (ToM) is often reported to decline with age. While early studies of ageing ToM produced mixed outcomes, the literature generally reports that ToM declines with ageing. However, many ageing ToM investigations employ paradigms that have sources of interference that are non-social in nature, comprising additional demands on executive function, which also declines with age. Consequently, these tasks may disproportionately disadvantage older adults. As a result, existing work may portray an inflated effect of age. To assess this, a meta-analysis was conducted. This showed that, while an overall age effect emerged, there was significant heterogeneity in the magnitude of the reported effects. The scale of the age effect varied and, in some cases diminished, based on paradigm type and sampling considerations, such as age and education, highlighting the problem of non-uniformity and non-specificity in ageing ToM. Next, using a novel paradigm, in-task manipulations which separated out executive demands that are relevant to ToM – outcome knowledge, self-other and attentional conflict – were examined to observe how these processes differentially affect older versus younger adults. Results showed that individual differences in attention and processing speed best explained the degree of conflict experienced through incongruent self-other perspectives. However, older adults were disproportionately affected by managing attentional conflict, and this source of interference was predicted by spatial working memory, highlighting that non-social demands can partially explain reduced ToM performance in ageing. Finally, fMRI was used to show differentiation in the neural response to social and non-social manipulations embedded within false belief tasks. Age-related differences in the recruitment of cortical regions, as informed by a ToM functional localizer, were also assessed. Overall, these findings bring greater clarity to our understanding of ToM in healthy ageing, disentangling classic experimental parameters to highlight which sources of interference best explain age-related difficulty in ToM reasoning.

Keywords: theory of mind; mentalizing; ageing; executive function; fMRI

بسم الله الرحمن الرحيم

For my little peanut – Haleema

Acknowledgements

I have been extraordinarily lucky to work under the mentorship of three world-class academics: Dr Charlotte Hartwright, Prof Klaus Kessler, and Dr Dan Shaw. Charlotte, thank you firstly for believing in me, and for your unwavering support throughout. You have taught me to be a principled scientist; that academic and scientific rigour comes before output or recognition. Klaus, your (characteristically German) mentorship both within and beyond academia has been transformational; thank you for backing me through the thick and thin. Dan, thank you for enduring the long Skype calls – it must have been so painful watching me fumble around aimlessly in FSL! Thank you for being so generous and supportive with your time.

I've been lucky to become friends with brilliant PhD students from Aston, past and present – there are too many to name here, but I would like to especially thank Nic Alexander, Adam Shephard, and Nathan Taylor, Lucie Thomas, Rachel Green, Cliona Kelly, and Nikki Ngombe. A special thanks also goes to an outstanding placement student – Dwayne May. Thank you also to everyone in office SW506, for the wonderful conversations, camaraderie, and moral support.

The Psychology Department, Aston Brain Centre, and ALIVE research facility also deserve a special mention. Special thanks also go to two brilliant radiographers at the ABC, Shaheen Lateef and Liz Squire. Thank you for putting up with my long scan sequences!

Thank you to my mum, dad, sisters Thameena, Najeerah, and (bhabi) Shahana, brother Mohibur, and of course my beautiful little niece, Mairah. Amma – without you, there simply would be no PhD. You set aside your own hopes and aspirations to make this all possible. To my mother-in-law, thank you for everything you have done, and especially the home-cooked Thai food – it made the writing process a lot less painful. To my Nana and Dada (رحمهما الله); no accomplishment comes close to repaying the sacrifices you made for us. Without your selflessness, this PhD would have remained an insurmountable, far-away dream.

Finally, two people who are the centres of my universe – my wife Margaret and to whom this thesis is dedicated, my daughter Haleema. Haleema, though the timing of your arrival was not ideal (raging global pandemic and PhD write-up year), your smile gives me the strength I need to tackle any and every challenge. Margaret, you have been my backbone and source of strength in the most testing of times. Your encouragement is the sole reason this thesis is now complete. I promise to never again be a 'student'.

List of Contents

Chapter One: General Introduction	9
Aims of the current thesis	31
Chapter Two	35
Introduction	38
Methods	53
Results	60
Discussion	81
Chapter Three	91
Introduction	93
Methods	97
Results	107
Discussion	125
Chapter Four	130
Introduction	133
Methods	155
Results	174
Discussion	196
Chapter Five: General Discussion	207
Conclusion	227
References	229
Appendices	253

List of Abbreviations

ACC	Accuracy
AC-PC	Anterior Commissure - Posterior Commissure
AQ	Autism Quotient
BET	Brain Extraction Tool
BOLD	Blood-Oxygen-Level Dependent
EF	Executive Function
EHI	Edinburgh Handedness Inventory
EV	Explanatory Variable
FB	False Belief
FEAT	FMRIB's Expert Analysis Tool
FLIRT	FMRIB's Linear Image Registration Tool
fMRI	Functional Magnetic Resonance Imaging
FMRIB	Oxford Centre For Functional MRI Of The Brain
FSL	FMRIB's Statistical Library
GLM	General Linear Model
GUI	Graphical User Interface
HRF	Haemodynamic Response Function
ICA	Independent Components Analysis
ISI	Inter-Stimulus-Interval
ITI	Inter-Trial Interval
m-ACE	Mini Addenbrooke's Cognitive Examination
MCFLIRT	Motion Correction FMRIB's Linear Image Registration Tool
MELODIC	Multivariate Exploratory Linear Optimized Decomposition into Independent Components
MMSE	Mini Mental State Examination
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
OA	Older Adult(s)
ROI	Region Of Interest
RT	Response/Reaction Time
SCF	Social Cognitive Function
SOCK	Spatially Organized Component Klassifikator
STS	Superior Temporal Sulcus
TB	True Belief
TFCE	Threshold-Free Cluster Enhancement
ToM	Theory of Mind
TPJ	Temporoparietal Junction
VOI	Volume Of Interest
WRAT	Wide Range Achievement Test
YA	Younger Adult(s)

List of Tables

Table 1.1. Summary of cognitive skills in older vs younger cohorts 19
Table 1.2. Comparison of models of cognitive ageing 23
Table 2.1. Examples of common theory of mind tasks 42
Table 2.2. Included studies and their respective participant and task characteristics. 63
Table 2.3. Statistical outcomes of sub-group analyses 73
Table 3.1. Summary of experimental factors and levels 99
Table 3.2. Summary of preregistered hypotheses 106
Table 3.3. Sample characteristics by groups: PT and EF 107
Table 3.4. Mixed ANOVA results for RT and error-rate 110
Table 3.5. Pairwise comparisons of RTs in YAs for OOC * SOC 116
Table 3.6. Multiple regression results for reaction time and error rate cost factors 121
Table 4.1. Theory of mind brain regions 139
Table 4.2. Summary of experimental factors and levels 157
Table 4.3. Tools and imaging applications 169
Table 4.4. Overview of contrasts in fMRI analysis 171
Table 4.5. Sample characteristics by group 174
Table 4.6. Trials removed per condition 175
Table 4.7. Mixed ANOVA on accuracy data 177
Table 4.8. Mixed ANOVA on RT data 178
Table 4.9. Pairwise comparisons comparing KOR and Age 180
Table 4.10. Correlations of experimental conditions (RT) and questionnaire measures 181
Table 4.11. Correlations of experimental conditions (Acc) and questionnaire measures 181
Table 4.12. fMRI regions in KOR contrast 186
Table 4.13. fMRI regions in K0S0 > K0S1 contrast 189
Table 4.14. fMRI regions in Belief > Photo contrast 192

List of Figures

Figure 1.1. Age differences in Fluid and Crystallised Intelligence 17
Figure 1.2. Scatterplots of age-related changes in brain volume 21
Figure 2.1. Growing trend of ToM-ageing publications 51
Figure 2.2. Search strategy and PRISMA flow diagram 61
Figure 2.3. Forest plot of studies included in meta-analysis 71
Figure 2.4. Funnel plot examining heterogeneity 80
Figure 3.1. False belief task diagram 100
Figure 3.2. Condition latencies and condition summaries 109
Figure 3.3. Schematic of key events in false belief vs. conflicting OOC trial 114
Figure 3.4. Estimated marginal means for error-rate and RT across YA and OA 117
Figure 3.5. Correlation matrix of all FB task conditions and EF measures 124
Figure 4.1. Mapping of different aspects of social cognition 140
Figure 4.2. Illustration of RMET 145
Figure 4.3. Factor structure of false belief task 157
Figure 4.4. False belief task outline and timings 160
Figure 4.5. Typical order of scan sequences and their respective parameters 165
Figure 4.6. Mean error-rates and RTs across experimental conditions 176
Figure 4.7. Interaction between Age Group and KOR 179
Figure 4.8. Accuracy and RT means as a function of Age Group 183
Figure 4.9. KOR 0 > KOR 1 effect axial slices 185
Figure 4.10. KOR 0 > KOR 1 effect 3D reconstruction 185
Figure 4.11. KOR & TPJ overlap 187
Figure 4.12. K0S0 > K0S1 effect axial slices 188
Figure 4.13. K0S0 > K0S1 effect 3D reconstruction 189
Figure 4.14. Belief > Photo effect axial slices 191
Figure 4.15. Belief > Photo effect 3D reconstruction 193
Figure 4.16. Mean parameter estimates of FB contrasts in ROIs 211
Figure 5.1. ATOMS framework 213
Figure 5.2. Percentage of population aged 60+ globally 216
Figure 5.3. Cognitive and affective theory of mind across lifespan 223
Figure 5.4. Trajectory of memory, processing speed, reasoning, and vocab against age 224

Chapter One: General Introduction

Humans are naturally driven towards cooperation and interaction, and this often requires a mutual or shared understanding of the complex, and typically muddled, world. As humans, and especially adults, we manage these social complexities with relative ease and thus typically do not expend much effort in appreciating the cognitive dexterities that underlie such processes. Central to social cooperation and understanding are the mentalistic notions of beliefs and desires, and how we use them in our understanding of others' mental and knowledge states. Inference, imputation, and representation of others' mental states – abilities commonly referred to as *mentalizing*, and collectively termed a *theory of mind* (ToM) – while not unique to humans (more of which later), are foundational to everyday, routine social functioning.

A functioning ToM has been shown to be associated with a wide range of real-world social outcomes in both children and adults. Peterson and Siegal (2002), for example, investigated mindreading abilities in 109 pre-school children that were divided into groups of 'popular' and 'rejected' based on mutual friendships and peer-popularity. Popular children, who had stable and reciprocal friendships, showed significantly greater understanding of false belief (FB) – representing others' mental states that are evidently untrue – relative to their less popular playground peers. Moreover, a reciprocal stable friendship accounted for variability in mindreading ability above and beyond the influence of age and verbal maturity. Similarly, in a prospective longitudinal study examining the link between friendlessness and ToM, Fink et al. (2015) studied 114 children over the course of two years and investigated ToM as a predictor of mutual friendship from ages 5 to 7 (Time 1 & 2, respectively). After controlling for language ability, Fink and colleagues reported that five-year-olds with a mutual friend fared significantly better compared to their friendless peers on a wide-ranging battery of ToM measures. In terms of a longitudinal effect, seven-year-olds who possessed no friends at either time of testing exhibited remarkably poor ToM at Time 1, even after controlling for group popularity, age, and language capacity.

Both Peterson and Siegal and Fink et al.'s findings mirror the key results of Slaughter et al.'s (2015) meta-analysis which examined ToM and peer popularity in early school years. Slaughter and colleagues, from their review of 20 studies comprising over 2000 children, reported that children with higher ToM performance also were more popular among their peers. Childhood friendships are important for longer-term psycho-social development as longitudinal studies of childhood friendships and peer popularity/acceptance have been shown to be related to psychological difficulties in later life. For example, friendless children have been shown to be statistically more likely to exhibit maladaptive psychological symptoms into adulthood, such as depression, anxiety, psychosomatic complaints, and lower self-esteem (Sakyi et al., 2015; Witvliet et al., 2010).

The corpus of ToM literature has typically focused on child and adolescent populations and therefore much less has been documented on the link between ToM and wider social outcomes in older age. An emerging literature, however, has shown the link between ToM in old age and broader social functioning. For example, older adults (OA) who demonstrate poorer ToM typically also show lower levels of social participation (e.g., being members of social clubs and societies, hobby groups, visiting friends and family, etc.; Bailey, Henry, & Von Hippel, 2008). Further, cognitive empathy[1] – a concept thought to be related to ToM – has been associated with measures of social

[1] There is a substantial divergence in what is thought to constitute a theory of mind, both in terms of a definition and in the methodologies used to investigate this phenomenon (e.g., see Quesque & Rossetti, 2020). This heterogeneity and lack of specificity is explored further through a meta-analysis in Chapter Two.

economic decision-making in advanced age (i.e., financial payoffs in the Ultimatum Game, Beadle et al., 2012).

Reduced social participation, which, as discussed above, is linked with poorer ToM, can be a serious concern in ageing due to its association with loneliness and social isolation. Research has shown that loneliness and isolation in the elderly are particular risk factors for all-cause morbidity and mortality (e.g., Pantell et al., 2013; Rico-Uribe et al., 2016), have been linked with obesity and high blood pressure (e.g., Freedman & Nicolle, 2020), and various mental health conditions, including depression (e.g., Smith & Victor, 2019). Further, the link between ToM and social-economic decision-making is of particular concern as fraud victimisation (being victims of irreversible financial fraud or exploitation) is most prevalent in the older age groups (see Shao et al., 2019). It is clear, then, that studying the precise nature of social cognitive function in ageing is important as there are potentially far-reaching consequences of ToM impairment in old age.

Origins of theory of mind

While early philosophical accounts of a 'representational mind' and 'beliefs about beliefs' were prevalent long before the contemporary understanding of mental state attribution (e.g., see Dennett, 1978; Pylyshyn, 1978), the formal coining of the term *theory of mind* is attributed to primatologists David Premack and Guy Woodruff. In their seminal 1978 investigation titled '*Does the chimpanzee have a theory of mind?'*, Premack and Woodruff defined ToM as the imputation of mental states to oneself and others, and used the terminology of 'theory' insofar as mental states are not directly observable but are understood through implicit and indirect concepts and methods of assessment (Premack & Woodruff, 1978). What constitutes a *mental* state was loosely outlined as intentions, beliefs, thinking, doubt, guessing, pretending, and liking.

In their study, Sarah, a 14-year-old African chimpanzee, was presented with 30-second videotaped scenarios of a human actor facing a set of four different challenges. These challenges were physical in nature, where the actor was depicted attempting to obtain particular objects that were out of his reach, or extricate himself from enclosed spaces. Following the videotaped problems, Sarah was shown a set of still photographs of potential solutions to the aforementioned problems. She was tasked with selecting the most appropriate solution to the problem from among the photographs. One example problem is where the human actor attempts to reach towards a bunch of bananas hanging from the ceiling, and the respective solution photograph showed an actor stepping onto a box. Sarah consistently selected the most appropriate solution to the respective problems, and this was interpreted by Premack and Woodruff as evidence of the animal showing an understanding of the actor's intentions, purpose, and desires (i.e., a ToM). This was the first experimental investigation into ToM.

However, in a commentary on Premack and Woodruff's seminal paper, Dennett (1978) questioned whether the chimpanzee genuinely represented the mental state of the actor, or, if more basic, lower-level problem-solving mechanisms could account for Sarah's success in resolving the scenarios. To specifically address this issue and provide robust evidence of mental state representation, Dennett – using the example of a children's Punch and Judy show – proposed a paradigm wherein the subject must hold in mind and reason about others' FBs. At its most basic form, FB understanding – which has subsequently been thought of as a litmus test for an operational ToM – requires one to manage the contrast between a perspective that is contrary to the true state of affairs (i.e., a FB) and what is evidently true (Wellman, 1993). Conceptually, FB is typically separated into first- (*John thinks that…*) and second-order (*John thinks that Sarah thinks…*) processes (Apperly, 2012; Astington, 1993; Sullivan, Zaitchik, & Tager-Flusberg, 1994).

On the basis of Dennett's proposition of requiring an understanding of 'falseness' in relation to one's own beliefs, Wimmer and Perner (1983) developed the now well-known object transfer task (also sometimes called the change-of-location task). In a visual sketch, participants are shown Character A placing an object into a location X and then leaving the scene. Then, in Character A's absence, Character B moves the object from X into location Y. As Character A was unaware of the change of location, participants had to assume that A still believed the object was in the original location. Participants are then asked where A will look for the object upon her return. Wimmer and Perner found that none of the 3-4-year old children in their study correctly pointed to location X; whereas 57% of 4-6-year olds and 86% of 6-9-year olds successfully resolved the task. The authors concluded that a functioning (though not complete) ToM develops around age 4-6 and that the higher proportion of errors experienced prior to this developmental milestone could reflect higher levels of cognitive egocentrism in pre-school-aged children. Wimmer and Perner's findings were the first to provide experimental evidence of a ToM in human subjects and their object–location protocol has become the template paradigm for subsequent and contemporary FB investigations.

Theory of mind disruption

As discussed above, children are typically successful on tests of ToM (in particular, false beliefs) around ages 4 – 6 (Wellman, 1992; Wimmer & Perner, 1983; for a recent review, see Slaughter, 2015). At this age, neurotypical children are able to understand the contrast in self-other beliefs and can effectively attribute and represent beliefs that are contrary to their own experience (Frith & Frith, 2005; Wellman & Liu, 2004). Disruption to the typical ToM trajectory can occur with the presence of particular developmental disorders and/or learning difficulties. In their seminal study, Baron-Cohen, Leslie, and Frith (1985) tested children with autism and Down's syndrome alongside a matched group of typically-developing controls. Based on Wimmer and Perner's (1983) original FB change-of-location task, Baron-Cohen and colleagues reported that children with autism showed significantly poorer mental state understanding compared to typically-developing children and children with Down's syndrome. The authors concluded that a disruption to ToM proficiency is independent of a global developmental and cognitive delay and is perhaps instead specific to autism spectrum disorders (for a detailed analysis of ToM in autism, see Tager-Flusberg, 2007). ToM continues to develop and mature through childhood and into adolescence, until stabilising in adulthood (see Dumontheil, Apperly, and Blakemore, 2010). Though, even in adulthood, where ToM competence seems to peak, individual differences in inhibitory control (Qureshi, Apperly, & Samson, 2010) and memory (Lin, Keysar, & Epley, 2010) have been shown to predict overall ToM performance. Disruptions to adult ToM can occur, however, as a result of a range of different psychiatric, neurological, and social disorders. These include, but are not limited to, schizophrenia (e.g., Kronbichler et al., 2017), depression and anxiety disorders (e.g., Bora & Berk, 2016), multiple sclerosis (e.g., Chalah et al., 2017), epilepsy (e.g., Morou et al., 2018), and in patients with acquired brain injury (e.g., Martín-Rodríguez & León-Carrión, 2010).

In healthy, non-pathological ageing, the extant behavioural and neuroscientific literature suggests an age-related impairment in ToM proficiency. However, the precise nature and magnitude of an age effect is still unclear as the earliest studies of ToM in ageing showed mixed outcomes. While more recent investigations have shown greater (but not complete) convergence with respect to an effect of age on social cognitive function (SCF), substantial theoretical and methodological heterogeneity still exist within the literature that precludes a definitive account being drawn on the relationship between cognitive ageing and ToM. For example, many investigations on ToM and ageing typically employ experimental protocols which require participants to manage non-social, fronto-executive demands – i.e., demands that older adults disproportionately find difficulty with. Thus, it is unclear from the existing knowledge base the extent to which non-social demands account for the reported deficit in ageing ToM, and how much of this reported ageing effect can be explained by genuine age-related modification in SCF. It is the purpose of this thesis, therefore, to disentangle an ageing ToM on a more granular level, to separate and identify the neurocognitive mechanisms that may explain variation in performance in older adults when reasoning about the beliefs of others. It is hoped, through the chapters/experiments in this thesis, a clearer understanding of how natural, healthy ageing modifies different features of ToM reasoning, both behaviourally and on a neural level.

Cognitive ageing

A thorough analysis of an age-related modification of ToM first requires explicating the wider cognitive degeneration that is present in healthy, advanced ageing. Cognitive abilities can be generally divided into specific domains including attention, memory, executive function, language, and visuospatial reasoning. Each of these domains can be further fractionated into smaller sub-processes (e.g., attention alerting, attention orienting, etc.) and the trajectory of each of these abilities differ in natural healthy ageing. Another important distinction is in the difference between fluid and crystallised abilities. Fluid aptitudes (or intelligence) require the on-line processing of novel information, to manage and resolve conflicting or distracting information or to reason about abstract and non-concrete problems. Crystallised abilities, on the other hand, rely on cultural or social learning; and can be typically thought of as aptitudes relating to general knowledge, vocabulary, and reasoning based on long-term, acquired information (Cattell & Horn, 1978; Horn & Cattell, 1967). Importantly, fluid and crystallised abilities have different trajectories with respect to ageing.

Fluid vs. crystallised intelligence

The most marked age-related change can be seen in *fluid* abilities (Bugg et al., 2006; Kievit et al., 2014; 2018; Salthouse, 2009; Salthouse et al., 1998). For example, Staff, Hogan, and Whalley (2014) examined 751 participants aged 62 – 83 in a longitudinal design where fluid intelligence was indexed using a variant of the Raven's Progressive Matrices (RPM). After standardising raw RPM scores to a mean test score of 100, Staff and colleagues reported that one additional year of life amounted to over .5 decrease in fluid intelligence.

In addition to longitudinal designs, cross-sectional investigations of fluid intelligence have also shown a reliable age-related deterioration, where, on average, individuals aged 60 score one standard deviation point below relative to 20-year-olds on measures including processing speed and efficiency (e.g., Salthouse, 2004; 2010; see Figure 1.1 below). Further, consistent linear decline in ageing has also been found on a wide range of measures of working memory (McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Salthouse, 1992) and attention (Craik & Salthouse, 2000). However, while there is a steady decline in fluid abilities with age, older adults typically show greater crystallised capacities relative to their younger counterparts. The greatest increment can be seen in abilities relating to language, typically assessed through tests of vocabulary and word pronunciation (Craik & Salthouse, 2011; Wechsler, 1997); though it should be noted that older adults experience greater difficulties in lexical retrieval (e.g., Segaert et al., 2018).

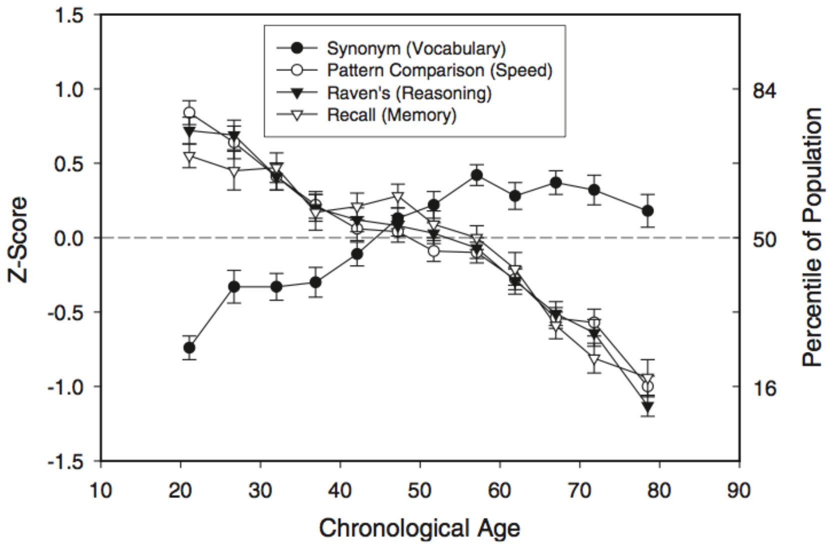


Figure 1.1. From Salthouse (2010): Age differences in Fluid and Crystallised Intelligence. Means (SD) of performance (Z-scores, percentile of population) across four tests as a function of age. Reproduced with permission.

Executive function

Executive functions (EFs; also known as executive control or cognitive control) encompass a set of top-down cognitive skills that drive attention, help set and maintain goals, resist temptation or primed responses, and preserve concentration when 'automatic' functioning would be ill-advised or insufficient (Miller & Cohen, 2001). Three core domains of EFs have been established in the literature (Miyake et al., 2000): 1) inhibition, which includes inhibitory control and interference control; 2) working memory

(WM), including working memory updating; and 3) cognitive flexibility, which includes abilities related to set shifting or task switching.

There is a substantial literature delineating a now well-established relationship between cognitive ageing and deterioration in all three components of EF. For example, older adults show increased difficulty, relative to their younger counterparts, at inhibiting both visual and auditory distractions (Alain & Woods, 1999; Gazzeley et al., 2005), and specifically show impairments in suppressing to-be-ignored distractors in attentional control tasks (Ashinoff, Geng, & Mevorach, 2019; Gazzeley et al., 2005; Mevorach et al., 2016). In stop-signal tasks, where participants are required to suppress or inhibit a pre-potent response, older participants again show poorer performance compared to younger adults (Hu et al., 2018). For a recent and comprehensive review of inhibitory control in ageing, see Rey-Mermet and Gade (2017). With respect to memory, there is an inverse relationship between age and the number of items that can be stored in WM (Craik & Salthouse, 2000; for meta-analysis, see Bopp & Verhaeghen, 2005). A well-known measure of WM is the digit span test which comprises forward and backward recall components. Choi et al. (2014), in their study of 784 community-dwelling, healthy older adults aged 60-90 years, reported a significant effect of age where older participants recalled fewer backward and forward digits overall. Finally, with respect to cognitive flexibility, Ferguson, Brunsdon and Bradford (2021) reported age-related impairment in a component of a task switching paradigm, where the authors reported a greater mixing (global task-shifting) cost in old age. Though, switching (local task-shifting) costs *decreased* in ageing. Further, in investigations measuring set-shifting through the Wisconsin Card Sorting Test, an age deficit is classically reported (e.g., Isingrini & Vazou, 1997; Raz et al., 1998). Taken together, though there is some heterogeneity across tasks, the literature reviewed thus far points toward a relatively robust association between ageing and weakening of executive control abilities. See Table 1.1 below for summary of changes in cognitive skills in ageing.

Table 1.1.
Summary of cognitive skills in older versus younger cohorts.

Cognitive Ability	Crystallised vs. Fluid	Age Effect
Attention	Fluid	Simple tasks: No difference Complex tasks: YA > OA
Executive function	Fluid	Inhibition: YA > OA Flexibility: YA > OA
Language	Mixed	Vocabulary: OA > YA Lexical retrieval: YA > OA
Processing speed	Fluid	YA > OA
Visuospatial	Fluid	YA > OA

Note: YA = Younger adults. OA = Older adults.

Neurocognitive ageing

A more fine-grained review of the ageing brain, with respect to both structural and functional changes, is provided in the Introduction of Chapter 4; accordingly, to avoid repetition and redundancy, I will outline only the broader models of cognitive and neurobiological ageing here.

The age-related impairment of cognitive abilities discussed previously can mostly be traced back to neurological modifications that are present in the brains of elderly individuals, which include alterations in neuroanatomy, functional and structural connectivity (FC and SC, respectively), and brain chemistry (for comprehensive reviews of neurocognitive ageing, see Cabeza et al., 2018; Grady, 2012; Harada, Love, & Triebel, 2013). First, with respect to neuroanatomical changes, ageing is associated with cerebral atrophy, particularly in the PFC, and studies have shown total grey matter volume begins to decline as early as within the third decade of life (Terry & Katzman, 2001). See Figure 1.2 below for illustration of age-related changes in brain volume, grey and white matter, and white matter lesions, separated by sex. On the specific link between atrophy and changes in cognitive function, Manard et al. (2016) tested young and older participants on a battery of neuropsychological measures, including inhibition, working memory, and set-shifting, and investigated changes to grey matter volume using voxel-based morphometry. Manard and colleagues reported a significant relationship between executive performance and age-related reduction in grey matter volume in anterior (frontal, insular, cingulate cortex) and posterior (temporal and parietal) brain regions. Though beyond the scope of this thesis, age-related atrophy and cortical thinning is thought to be associated with a proliferation of beta-amyloid protein and its association with neuronal death in ageing (e.g., see Harada, Love, & Triebel, 2013).

Changes in SC and FC may in part also drive behavioural differences between young and older participants. Briefly, FC is an index of the temporal linking of discrete brain regions whereas SC describes the integrity and presence of white matter tractography that interconnects separate regions (Uddin, 2013). Zonneveld et al. (2019) examined resting-state FC in 2878 healthy OA in a Dutch population-based study. Resting-state scan sequences derive data participants in a state of 'rest'; where participants are typically instructed to 'do nothing'. Resting state data is useful therefore in quantifying a baseline measure of intrinsic (i.e., task-free) activity, and is commonly used in the study of FC. Zonneveld and colleagues reported significantly lower FC within the anterior default mode network, ventral attention, and sensorimotor networks in older participants. Though, it useful to consider that FC was higher within the visual network in the older group. Linking FC and SC in ageing to cognitive performance, Fjell et al. (2017) reported that both SC and FC changes were related to longitudinal reductions in executive function.

Finally, there is a well-established literature showing age-related changes in brain function or 'activity' across a number of cognitive tasks. Activity, in the context of fMRI, is typically indexed as fluctuation in the blood-oxygen-level-dependent (BOLD)

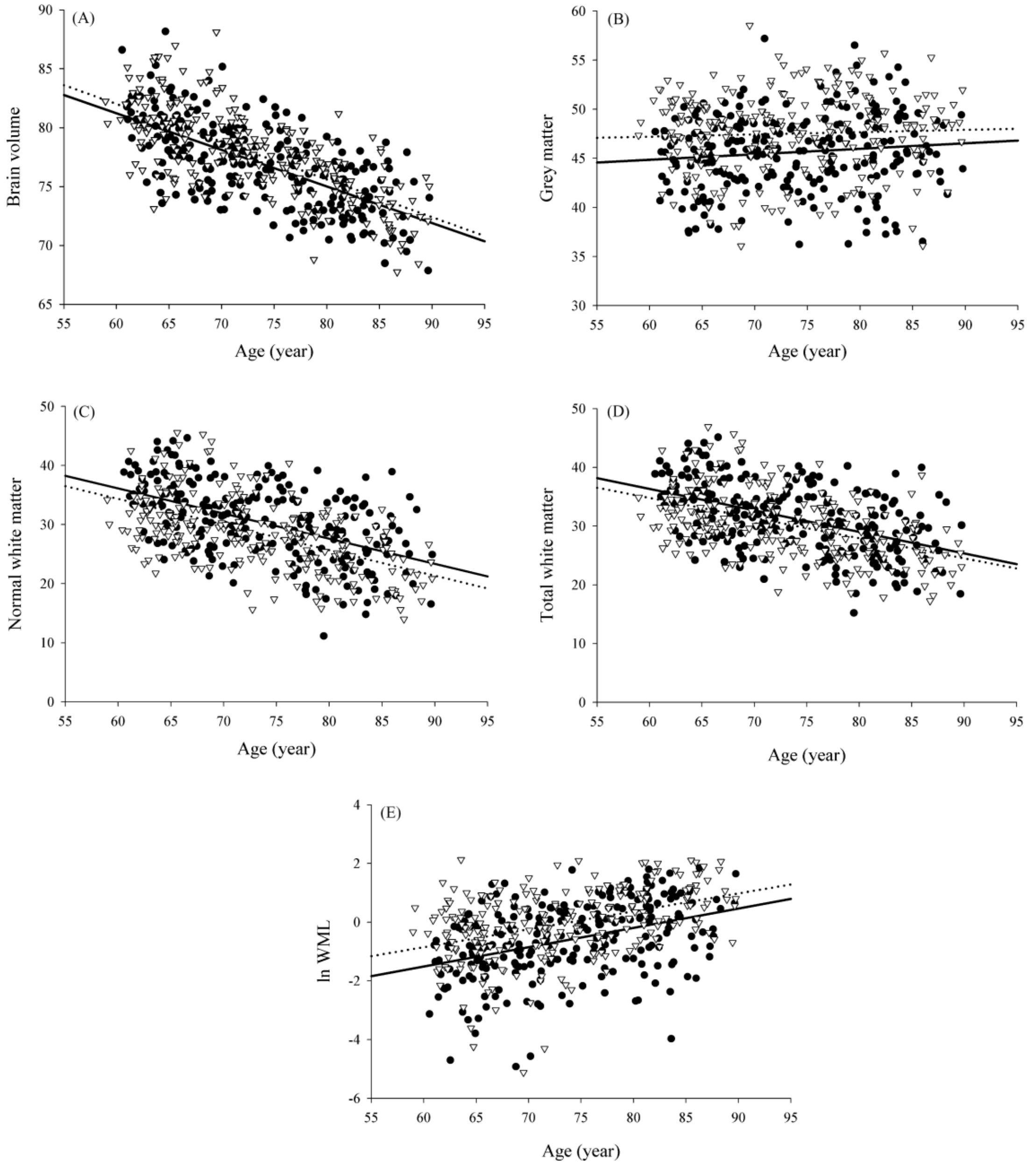


Figure 1.2. Scatterplots of age-related changes in brain volume, separated by sex. Reused with permission from Ikram et al. (2008): 10.1016/j.neurobiolaging.2006.12.012 (A) Brain volume; (B) grey matter; (C) normal white matter; (D) total white matter; (E) white matter lesions (natural log transformed). Volumes are expressed as percentage of intra-cranial volume. Women: open triangles and dotted line. Men: closed circles and solid line.

signal in response to stimuli. A more comprehensive analysis of the BOLD signal in ageing is provided in Chapter 4. Although fMRI studies of ageing have shown that peak stimulus-related BOLD responses are comparable in young and older subjects (e.g., Aizenstein et al., 2004; Huettel, Singerman, & McCarthy, 2001), there are task-based differences between young and old age groups in *where* the neural activity is concentrated. There are two primary conceptual devices that can be used to explain differential recruitment and activity between young and older participants: *compensation* and *dedifferentiation* (for a review of these processes, see Grady, 2012).

Compensation describes the phenomenon wherein OA sometimes engage particular brain areas, particularly within the frontal lobes, above the level seen in YA to compensate for weakened function somewhere else in the brain (Grady et al., 1994). For example, older adults have been shown to display more activity in the PFC during memory tasks compared to their younger counterparts (e.g., Madden et al., 1999). This has been interpreted as compensation for lower activation in other parts of the brain, particularly the visual processing regions (Grady et al., 1994). The preponderance of frontal activity in ageing, in comparison to poster regions, is known as the *posterior–anterior shift with ageing* (or PASA). See Table 1.2 below for a summary of models of neurocognitive ageing. Dedifferentiation, which is not mutually exclusive to compensation, describes the more diffuse, bilateral patterns of activity seen in OA on tasks where YA show lateralised and circumscribed – or asymmetric – activity (Madden et al., 1999). For example, Dennis and Cabeza (2011), in an ageing fMRI study of implicit and explicit memory, found that YA showed increased hippocampal activity during explicit learning and increased recruitment of the striatum during implicit learning. OA, however, showed less specific activity during these two conditions, and

Table 1.2.
Comparison of models of neuro-cognitive ageing

Model Acronym	Model Name	Key Publication(s)	Key Postulates of Model
CRUNCH	Compensation-Related Utilization of Neural Circuits Hypothesis	Reuter-Lorenz & Cappell (2008)	The model posits that older adults are generally more likely to show increased activation than younger adults during tasks that have low-level demands, even when performance between groups is comparable. When task-related demands rise, older adults reach capacity for compensatory recruitment, and further compensation becomes no longer possible. Accordingly, during particularly challenging tasks, activation levels of older adults may drop below those of younger adults. The point of inflection on the task demand axis where activity reaches its optimal (peak) level and begins to decrease is known as the *crunch* point. This point is specific to individuals and could be shifted by exercise, cognitive training, neural damage, sleep deprivation, or genetic predisposition.
HAROLD	Hemispheric Asymmetry Reduction in Older Adults	Cabeza (2002)	The HAROLD model suggests that compared to younger adults, who show unilateral patterns of activity, older adults show greater bilateral activity, especially in the prefrontal cortex. For example, during verbal and spatial working memory tasks, older adults exhibited bilateral prefrontal activity compared to younger adults who showed left-lateralised recruitment during verbal and right-lateralised recruitment during spatial working memory tasks (Reuter-Lorenz et al., 2000). The increased laterality is thought to drive either compensatory processes or dedifferentiation; these mechanisms are not considered to be mutually exclusive under this Model.
PASA	Posterior-Anterior Shift in Ag[e]ing	Davis et al. (2008); Dennis & Cabeza (2008)	The PASA model suggests that older adults show less activation of posterior regions (e.g., occipital lobe) and greater activation of frontal regions, compared to younger adults. The preponderance of frontal activity in older adults may be present even in the absence of behavioural age differences in task performance. The PASA model favours the compensation hypothesis, where increased recruitment of frontal regions in older adults is thought to offset under-activity/impairment elsewhere in the brain.

STAC	Scaffolding Theory of Aging and Cognition	Park & Reuter-Lorenz (2009)	The STAC theory posits that older adults recruit compensatory neural scaffolding as an adaptive response to declining brain structure and brain function. Those older adults with less neural degeneration, greater ability to recruit compensatory scaffolds, or both will have better cognitive performance than older adults who are unable to use scaffolding to compensate for their neural declines. Scaffolding involves the recruitment of additional, complementary neural resources to sustain dedicated circuitry to preserve cognitive performance.
STAC-r	Scaffolding Theory of Aging and Cognition - revised	Reuter-Lorenz & Park (2014)	STAC-r is a revision of the original STAC model. In addition to incorporating all aspects of STAC, the revised model also includes flexibility to account for longitudinal change in cognition and individual differences in life experiences, which can be categorised as either neural enrichment (e.g., education, physical fitness) or neural depletion (e.g., depression, vascular disease) factors. These factors are thought to impact brain structure and function, and the potential for compensatory scaffolding, which collectively predict the level of cognition and rate of cognitive change. STAC-r is the broadest model in this comparison, as it factors in flexibility to account for fluctuation in cognitive change over *time*.

showed comparable activity in both aforementioned regions during implicit and explicit memory conditions. Overall, natural healthy ageing brings about material changes in a wide range of cognitive skills; and the modification of these skills is sometimes associated with altered or impaired patterns of neural activation. This background to cognitive ageing is important to consider when examining an age-related weakening of ToM.

Interplay between executive control and ToM proficiency

At the broadest level, ToM comprises many different constructs and each construct has a multitude of unique methods and experimental protocols. For a review of the different facets of ToM, see Quesque and Rossetti (2020). The ToM component that is the focus of this discussion, and indeed the wider dissertation, is FB understanding. FB is most commonly (though not exclusively) assessed using the previously-discussed object-transfer paradigm, and the three measures of EF – inhibition, working memory updating, and attention shifting – can be mapped somewhat directly onto the subcomponents of the object-transfer task. For example, a typical representation of the task includes multiple agents, each with their own mental state and multiple – and often competing – locations. In order to represent the mental state of another agent, a participant must suppress their own knowledge of reality and this ostensibly draws on mechanisms of inhibition or inhibitory control. The shifting of attention from one spatial location to another – e.g., the *basket* and the *box* – is thought to require attention shifting (cognitive flexibility) skills while keeping in mind the contents of each location and the subsequent change of the target location is thought to require working memory updating. More recent FB paradigms comprise more complex problems where more than two locations are presented (e.g., three locations in Wysocka et al., 2020) or have more intricate visuo-spatial demands (e.g., the 'sandbox task' in Bernstein, Thornton, and Sommerville, 2011), and therefore perhaps draw on these executive abilities more deeply.

From a developmental perspective, it is important to consider causality or 'direction': does emerging EF drive developments in ToM, or is the development of ToM central for an advancement in EF capacities? Hughes and Ensor (2007) addressed this issue directly in their study of 122 socially diverse children tested at ages 2, 3, and 4. The authors examined the predictive relationship between ToM and EF using partial correlations and hierarchical regression models, which included factors of verbal ability and social disadvantage in addition to tests of ToM and EF. Hughes and Ensor concluded that the strongest predictive relationship supported the view that EF abilities facilitated children's performance on ToM tasks. Indeed, numerous longitudinal and cross-sectional developmental studies have examined the specific ToM-EF relationship and have reported EF as a precondition to an operational ToM (e.g., Austin, Groppe, & Elsner, 2014; Carlson, Moses, & Breton, 2002; Hughes, 1998; Marcovitch et al., 2015; Perner & Lang, 2000).

However, only a limited number of studies have attempted to delineate the relationship between EF and ToM in non-clinical, adult populations. In one study, researchers analysed the predictors of cognitive and affective ToM in young and older adults separately (Fischer, O'Rourke, & Thornton, 2017). Digit span, Colour-Word Stroop, and Letter-Number Sequencing measures were performed in addition to semantic and episodic memory and pulse pressure (the difference between systolic and diastolic blood pressure values, in mmHg). Path analyses showed that EF, semantic memory, and pulse pressure predicted cognitive ToM in young adults, whereas semantic memory was the only predictor for affective ToM (Fischer, O'Rourke, & Thornton, 2017). It is difficult to establish precisely which measure of EF (e.g., inhibition, working memory, or attention) was most predictive of young adult ToM as the authors did not test these predictors individually. Bull, Phillips, and Conway (2008) on the other hand used dual-task manipulations of EFs (inhibition, updating, and switching) to examine the influence of these cognitive functions in mental and non-mental state tasks in 150 young participants (age range 16-31). The authors reported ToM task costs when concurrently performed with an inhibition task (but not updating or switching), suggesting a specific link between inhibitory control and mental state reasoning. However, the evidence to date linking specific and individual cognitive control functions with specific manipulations of ToM tasks in healthy younger adults is limited.

In advanced age, executive abilities have been shown to predict ToM performance. Phillips et al. (2011) studied 129 adults aged 18 – 86 on visual and verbal[2] FB paradigms and found that working memory updating partially mediated age-based differences in FB reasoning. Likewise, Nolaker et al. (2018) reported that, above and beyond the effect of age, a combination of measures of EF explained individual differences in performance on a composite of three naturalistic tests of ToM. Finally, in their study of 27 young and 20 older adults matched on crystallised intelligence, Rakoczy, Harder-Kasten, and Sturm (2012) examined the influence of domain-general abilities in old-age ToM decline. General cognitive skills were measured using the Trail Making Test, the Stroop task, a processing speed measure, and the day-night task (an index of interference control). SCF was measured using two ToM tasks: a German translation of Happé et al's (1998) Strange Stories paradigm and Sullivan and Ruffman's

[2] Theory of mind (false belief) paradigms are typically constructed as either verbal or non-verbal tasks. Verbal tasks predominantly present stimuli textually (i.e., vignettes) whereas non-verbal tasks usually present stimuli pictorially (i.e., cartoon/comic strips or video-taped scenarios). A more in-depth discussion on verbal/non-verbal tasks can be found in Chapter 2.

(2004) video task. Overall, OA performed worse on both measures of ToM (more of which later) but importantly the age-related deficits were partly explained by domain-general decays in processing speed and EF (Rakoczy, Harder-Kasten, & Sturm, 2012). Though the effect of age was not completely abolished, these results indicate that a potential effect of age is in part attributable to wider cognitive functions; which also happen to decline with age.

Theory of mind in old age

A comprehensive analysis of ageing ToM is provided in the empirical chapters of this thesis; therefore, I will only briefly discuss the broader literature here. In their meta-analysis of 23 studies comprising over 1400 participants, Henry et al. (2013) examined the effect of age on ToM across a range of tasks, modalities (visual, verbal, static, dynamic, etc.), and domains (cognitive, affective, or mixed). Overall, collapsed over all task types, OA were reported to perform more poorly than YA. The effect of age was not homogenous, however; there was variability in the magnitude of the age effect depending on the specific sub-group analysis, suggesting that task and sampling parameters might in part explain some of the difference in young vs. older adult ToM performance. The proposed effect of age has been further demonstrated in more recent empirical investigations using a range of different paradigms, including animated cartoon films (e.g., Baksh et al., 2018), naturalistic social movies (e.g., Lecce et al., 2019), first-person virtual reality tasks (e.g., Grainger et al., 2020), in addition to traditional vignette-based paradigms (e.g., Hughes et al., 2019; Rakoczy et al., 2018).

It is worth noting, however, the evidence is not entirely homogenous respect to an age effect. Indeed, the earliest study of the effect of ageing on SCF reported *enhanced* abilities in older subjects (Happé, Winner, & Brownell, 1998). However, Happé and colleagues noted that the young and old groups were not IQ-matched, and it seems probable that selection bias may explain this effect. Even in studies where age differences are present, there is substantial heterogeneity in the magnitude of effects, as the effect seems to vary depending on a range of factors relating to sampling (e.g., age and education) and task design (modality, domain, and preferred outcome measure). The nature and magnitude of an age effect is explored in further detail in a meta-analysis in Chapter 2. The non-uniformity can perhaps be explained by qualitative differences in task difficulty. In addition to theoretical heterogeneity in the traditional understanding of what ToM actually constitutes, there is substantial divergence in task design, with particular tasks placing greater demands on non-social executive processing (for a recent analysis of ToM tasks, see Quesque & Rossetti, 2020). For example, where some naturalistic tasks are rich in detail and provide comprehensive and individualised backstories to the narratives accompanying protagonists, animated shapes tasks are devoid of this social 'richness' and present only moving shape forms that perhaps have lower demands on particular EFs, such as working memory. Further examples of this inconsistency can be seen even within the *same* task type. Consider Wimmer and Perner's (1983) classic object-location paradigm and a more recent variant by Bernstein, Thornton, and Sommerville (2011). As discussed previously, the conventional object-location (also known as change-of-location) task typically requires participants to infer the beliefs of a character who holds a FB regarding the location of a target object, which is contrasted with the participant's updated and privileged belief of the scene. The participant is critically asked to indicate which location the protagonist will search upon her return in order to retrieve the target object. In Bernstein, Thornton, and Sommerville's (2011) sandbox task, a similar problem is offered to the participant, but instead of indicating which two-dimensional location the protagonist will search, the participant here is asked to manually point to a location within a physical sandpit where the protagonist is likely to search. The physical distance between the correct location and the participant's indicated position – as measured by a ruler – is taken as the outcome variable. Ostensibly, both paradigms are purported to test FB understanding, contain identical numbers of protagonists and 'correct' locations, and are non-verbal in their format in that the stimuli are visually presented, and thus can be argued to have similar processing demands. Therefore, any age effect resulting from these paradigms might be considered comparable. However, it is at least arguable that there are greater visuo-spatial/motor processing demands – which are known to diminish with age (e.g., Bendayan et al., 2017) – in Bernstein and colleagues' sandbox task as it requires participants to visually orient and estimate a physical, three-dimensional location. These differences in task design may introduce additional non-social processing that disproportionately disadvantage OA, and in turn inflate any genuine age-related differences in SCF. It is important, therefore, to examine ageing ToM using carefully-constructed tasks wherein social and non-social manipulations are systematically varied to highlight precise sources of interference, and delineate which of these processes are differentially affected by age.

Models of cognitive ageing and age-related theory of mind decline

As discussed previously, there are a number of different models which explain the age-related deterioration of cognitive function both on a behavioural and neural level. The models are mostly iterative with shared underlying concepts and mechanisms; however, for the current work—that is, ToM reasoning in healthy ageing—the STAC-r model (Reuter-Lorenz & Park, 2014) represents the most compatible account. An extension of the original Scaffolding Theory of Aging Cognition model (STAC; Park & Reuter-Lorenz, 2009), the revised STAC model is flexible in that it accommodates life-course variables such as education, cognitive and physical training, disease, genetics, and so-on, and thus allows for a more fine-grained analysis of individual differences in ageing. While there is a general downward trajectory in most (but not all) cognitive functions in ageing, as discussed previously, the speed and extent of decline varies greatly depending on a range of individual factors (see Harada, Love, & Triebel, 2014). Similarly, age-associated deterioration of ToM is not uniform (see Chapter 2 for a more comprehensive discussion on non-uniformity in ageing-ToM). For example, individual differences in EF have been shown to predict ToM performance in older age (e.g., Rahman et al., 2021). Moreover, studies of ToM training, where authors have aimed to promote social-cognitive abilities in older adults, have also found that ToM is indeed sensitive to practise and learning (see Cavallini et al., 2015; Rosi et al., 2016). That ToM ability in ageing varies as a result of individual factors and is sensitive to training effects, suggests that STAC-r is perhaps the most compatible model of neuro-cognitive ageing in the context of social cognition.

Aims of the current thesis

At the broadest level, the aims of this thesis were to clarify and bring greater specificity to a potential effect of age on ToM functioning. As discussed above, there is substantial variability in the magnitude of age effects in the literature, with a number of studies showing comparable performance between young and older groups. This inconsistency is in part driven by non-standardised, non-uniform tests of ToM where processing demands or cognitive load varies as a function of task type. The variation in processing demands, especially those processes that rely on capacities known to diminish with old age (e.g., EF), may partially explain the discrepancy in effects both within and between age groups. The first aim of this thesis, then, was to elucidate the precise methodological parameters (e.g., task design, sampling, etc.) which might differentially contribute to the oft-reported effect of age. Systematically probing an age effect across a number of theoretically-driven task and sampling considerations will provide a more nuanced and granular account of age-associated changes in ToM function. This aim is realised primarily through a meta-analysis in Chapter 2.

The second general aim of the thesis was to examine precisely the sources of difficulty in reasoning about others' mental states, and how the sources of interference may differentially change with age. As alluded to previously, social and non-social demands are typically confounded in the social cognition literature. While this is perhaps not a critical issue in ToM studies with younger participants employing within-subjects designs, the matter of differentiating processes that are fundamental to belief reasoning from those that are driven by more domain-general abilities is important when examining an age effect. As discussed above, FB tasks contrast the participant's true and privileged belief with the protagonist's outdated and superseded belief; and the competition arising from keeping in mind these two conflicting belief states is a source of difficulty or interference. At the same time, FB tasks, through methodological confounds as discussed above, such as managing multiple and competing locations, introduce interference that is non-social in nature. This particular aim, then, focused on delineating social and non-social parameters to evaluate which cognitive components associated with belief-reasoning explain age-specific deficits in performance. This particular aim was addressed through two studies – Chapters 3 and 4 – as outlined below.

Thesis structure

In addition to this current Introduction (Chapter 1), there are four subsequent chapters in this thesis. Chapter 2 comprises a meta-analysis of behavioural studies investigating age-related decline in ToM. Chapter 3 is an empirical investigation of the precise sources of interference in FB reasoning, and how these sources of interference differentially impact performance between young and older participants. Chapter 4 is a functional magnetic resonance imaging (fMRI) study of FB reasoning in older age, where, using a modified version of the task employed in Chapter 3, we attempted to delineate a neurocognitive account of ToM reasoning in healthy ageing. Finally, Chapter 5 comprises a general discussion which includes limitations, future directions, and a conclusion. Each of the three empirical chapters – 2, 3, and 4 – are outlined briefly below.

Chapter 2. What follows in Chapter 2 is a meta-analysis of behavioural investigations of age-related ToM decline, where we specifically examined a young vs. older comparison across a range of theoretically-driven sampling and ToM task-related variables. Previous meta-analyses have shown an overall age-related decline in ToM abilities. We extended the literature by examining the heterogeneity in the effect of age across two novel sampling considerations, explicitly probing the reported age effect as a function of OA' mean age and level of education. We additionally examined differences in ToM performance in domain (cognitive vs. affective) and task modality (verbal vs. no-verbal).

Chapter 3. In the first empirical investigation, we broadly studied if FB reasoning is effortful for OA beyond the non-social cognitive demands of classic ToM investigations. We assessed older and younger adults' ToM on a novel FB paradigm, and examined how age interacted with three theoretically relevant sources of conflict within ToM: competing Self-Other perspectives; competing cued locations, and outcome knowledge (or a knowledge of reality). We also used measures of EF to assess the relationship between successful belief reasoning and executive abilities, and examined if this relationship changed with age.

Chapter 4. Building on Chapter 3, we again tested young and older adults on an amended version of our novel FB paradigm, while administering fMRI. In an effort to again examine the precise nature of difficulty in ToM reasoning across young and older adults, we manipulated whether the participant had a privileged knowledge of reality and varied the congruence between Self and Other beliefs. We also administered Saxe and Kanwisher's (2003) classic ToM localizer to inform our analyses of the fMRI data. Further, for the first time, we included in our neuroimaging protocol an index of neurovascular coupling to account for age-related differences in neurophysiology. In addition to probing behavioural differences, we aimed to delineate between our task manipulations on a neural level, and aimed to show differential recruitment of established regions within the classic mentalizing network (as informed by the localizer) between young and older participants.

Thesis format. The chapters in this thesis were written in a 'publication format', where each chapter was constructed as a self-contained manuscript. To this effect, naturally, there is a degree of redundancy and repetition, especially with respect to the introductory segments across the chapters. Lastly, while this work is my own, I frequently (and preferably) use the term 'we' to better acknowledge the collaborative nature of studying for a PhD, which has benefitted and taken from the expertise of many academics and scholars.

Chapter Two: Meta-analysis of ToM in ageing

Meta-analysis of the nature and magnitude of age-related change in theory of mind.

Chapter précis and background

Early studies of theory of mind (ToM) in ageing produced mixed results, with some studies suggesting preserved or even enhanced social cognitive function in advanced age. More recently, the literature has converged (though not completely) to suggest that natural, healthy ageing brings about deterioration of ToM irrespective of task domain and modality. Indeed, meta-analyses have shown an age-related deficit across a wide range of ToM protocols, suggesting a stable and robust effect of age. However, these quantitative reviews have frequently reported significant heterogeneity in effect sizes when examining performance between young and older adults. This variability is perhaps explained, at least in part, by substantial theoretical and methodological heterogeneity in the extant literature, with considerable variability in conceptualising what ToM actually is, what ToM tasks actually measure, and the cognitive social and non-social demands that are unique to each ToM paradigm. Against this backdrop, here in Chapter 2, I examined the effect of age through a meta-analysis comparing young vs. older adults' performance across a range of different theoretically-motivated analyses. In an effort to probe the heterogeneity in effect sizes, we specifically examined the effect of age as a function of task type, domain (cognitive, affective), and modality (verbal, non-verbal). Crucially, we extended the literature base by examining an age effect across two further considerations with respect to sampling demographics: age and education level of older adults.

Overall, in line with previous meta-analyses, we found a general effect of age, with poorer ToM performance observed in older compared to younger adults. However, our analysis showed significant heterogeneity in observed effect sizes. The predetermined sub-group analyses showed an effect of age across domain and modality categories; though there was again significant heterogeneity even within sub-group analyses, with some particular categories showing no effect of age. Likewise, the overall effect was also seen within the categories of education and age, though, as before, the magnitude of the effect varied substantially. Finally, the education and age sub-groups were combined to examine if there was interaction behind these two variables. As expected, the effect again showed specificity according to sub-group, with older and less educated older adults showing the worst ToM performance.

The use of sampling variables is novel to meta-analyses of ToM in ageing, and, at the time of writing, has not been tested prior to this current investigation. While we replicated an overall effect of age, we showed that the reported age effect is closely dependent on the nature of the tasks being employed and the sampling make-up of the population being tested. Indeed, in some sub-analyses, there was no effect of age. In sum, it is likely that the classically reported effect of age has been somewhat inflated due to differential task demands and a lack of consensus regarding testing protocols and sampling considerations. These findings have important implications for future investigations of ToM and ageing and we recommend greater methodological and theoretical standardisation, which are further detailed herein.

Introduction

The social cognitive ability to impute, infer, and respond to others' mental states – i.e., having a theory of mind (ToM) – is crucial to navigate the social world (Frith & Frith, 2005). Disruption or difficulties in reasoning about the mental states of others, especially appreciating a Self-Other distinction, can have profound social consequences that may diminish one's quality of life (e.g., see Bailey, Henry, & von Hippel, 2008). Namely, individual differences in adult ToM have been linked with peer acceptance, peer rejection, and social anxiety (Ronchi, Banerjee, & Lecce, 2019).

In humans, an intact, operational ToM develops around age 4 – 5 and remains relatively uninterrupted until older adulthood (Wellman, 1992; Wellman, Cross & Watson, 2001), where age-related deficits have been reported in a range of ToM tasks including false belief reasoning (e.g., Phillips et al., 2011), detection of social faux pas (e.g., Bottirolli et al., 2016), and detection of double-bluffs, mistakes, and white lies (e.g., Maylor et al., 2002). However, there is inconsistency in the literature regarding the nature and magnitude of any age effect, and indeed some studies have reported no difference or even improved ToM in old age (e.g., Happé et al., 1998). This disparity may be explained by substantial heterogeneity across ToM protocols and the wider non-social cognitive demands within social cognitive tasks. The lack of specificity in classic ToM measures may have resulted in an under-estimation of ToM in healthy older adults: differences in performance between younger and older adults may be, at least in part, explained by methodological and sampling properties. To address the lack of clarity in the ageing ToM literature, this review will assess the nature and magnitude of an effect of age in ToM, using theoretically-driven sub-analyses to interrogate the potential age effect across various task and sampling considerations. This will provide a better understanding of what ToM looks like in older adults, when taking into account methodological considerations.

ToM, at the broadest level, refers to both cognitive and affective components (or *modalities*). Where cognitive ToM involves imputing and representing intentions and knowledge states, affective ToM refers to the appreciation and attribution of others' feelings and emotional states (Kalbe et al., 2010). Early accounts of ageing ToM, and indeed ToM more generally, however, seldom differentiate cognitive from affective components and typically confounded one capacity with the other (for a discussion, see Apperly, 2010). This is potentially problematic as the two modalities have been shown to be distinct both on a behavioural and neural level (Sebastian et al., 2012; Shamay-Tsoory & Peretz, 2007); and have different developmental trajectories in advanced age (Bottiroli et al., 2016; Fischer, O'Rourke, & Thornton, 2017). Specifically, where cognitive ToM has been reported to decline with age, affective ToM appears to be preserved (for a recent investigation on the differentiation of the two modalities of ToM in healthy ageing, see Ruitenberg, Santens, & Notebaert, 2020). With respect to ageing ToM, then, this distinction between the two modalities should be considered with respect to the nature and magnitude of an age-related decay in ToM.

At the broadest level, this heterogeneity is manifested in the proliferation of non-specific terminology that is used to loosely describe ToM and its associated processes. Terms such as *mentalizing*, *mindreading*, *empathy*, *cognitive empathy*, *perspective-taking*, *belief reasoning, mental state attribution,* among others, ostensibly all refer to a single underlying concept: *the imputation or attribution of mental states to oneself and others* (Apperly, 2012; Quesque & Rossetti, 2020). In their analysis of tasks purported to test ToM, Quesque and Rossetti (2020) identified 23 distinct paradigms. They categorised each according to whether there was a necessity to represent mental states (which the authors termed *Mentalizing criterion*), and if the task distinguished between self and other mental states (*Nonmerging criterion*). The authors argue that these criteria form the foundation of social inference as they involve representing mental states that differ from what is ostensibly true in reality, and being cognisant of and understanding the distinction between one's own beliefs and the beliefs of others (Apperly, 2010; 2012). Critically, in their review, Quesque and Rossetti identified that many paradigms – although purported to be 'ToM tasks' – did not comprise mentalizing nor self-other distinction elements, with several tasks lacking both mechanisms. These tasks, that are devoid of mental state representation, may in fact tap lower-level basic cognitive processes as opposed to social cognitive function; and according to the authors, tasks that do not satisfy the Mentalizing and Non-merging criteria "[S]hould no longer be discussed as a measure of theory of mind" (Quesque & Rossetti, 2020, pp. 386). As can be seen from Table 2.1 below, the variability in age effects across studies is perhaps reflective of the theoretical underpinning of the different task types. Some tasks are qualitatively more *difficult* to resolve than others as they tap into divergent facets of ToM and wider social cognition. Where some tasks require the detection of awkward or inappropriate human behaviour in complex social scenarios (e.g., the Faux Pas test), others require discerning the intentional state of non-human, inanimate shape forms (e.g., Castelli et al., 2002). These tasks may pose substantially different demands on supporting systems for language and executive functions, which in turn may drive variability in performance in older age, and modify the reported difference in SCF between young and old.

On a task level, the inconsistency of age effects may reflect the substantial variation in incidental and non-social cognitive demands embedded within ToM tasks. For example, consider Baron-Cohen et al.'s (1985) well-known FB change-of-location task: the 'Sally-

Ann' paradigm. Based on the first experimental investigation of children's FB by Wimmer & Perner (1983), participants are presented with two characters – Sally and Ann – in a static picture-based paradigm. Sally, who is playing with a ball, places her ball in a basket and leaves the scene. While Sally is away, Ann removes the ball from its location and places it inside a box (unbeknown to Sally). In the final part, Sally returns to the scene and the participant is asked where Sally will look for her ball. To successfully resolve the task, participants must suppress their own knowledge of reality (or true belief; *the ball is in the box*) to keep in mind and represent a false belief (*the ball is [still] in the basket*). Often, there is more than one location to keep track of. Further, where some studies require mentalizing about one single character, others present multiple characters each with their own distinct mental states. This potentially drives further variability as there may be an increased burden or cognitive load when keeping in mind and reasoning about multiple mental states simultaneously compared to attributing beliefs to a single character. These processes manifestly rely on working memory (multiple locations, multiple agents), attention or attention switching (shifting attention between candidate locations), and inhibitory control (suppression of one's own [often privileged] belief). Indeed, studies have shown that EFs, and in particular inhibitory control, are strongly related to false belief reasoning even when age and general intelligence are controlled for (Carlson et al., 2002; Kloo & Perner, 2003).

Table 2.1.
Examples of common theory of mind tasks and their respective properties.

Task type	Paradigm	Example	Domain	Modality
Detection of socially awkward or inappropriate utterances	Participants must detect and respond to faux pas set within social scenarios. A faux pas, which typically has negative social consequences, happens when a speaker utters something that is inappropriate or awkward, without considering if the listener is aware of the true state of affairs.	Faux Pas Test: Baron-Cohen et al. (1999).	Cognitive / affective	Verbal
Inference of mental states of protagonists in short fictional stories.	Original Strange Stories paradigm comprised 24 short passages, which were of three types: ToM, control, and jumbled text. ToM stories centred on double bluffs, mistakes, persuasions, and white lies. Each passage was followed by a test question, and for ToM trials, this involved inferring story characters' intentions. Control stories and test questions were based on physical (not mental) processes, i.e., physical causation.	Strange Stories: Happé et al. (1994).	Cognitive	Verbal
Inferring mental-emotional states from photographs	Participants are presented with 25 photographs of only the eye-region of faces and are asked to match words and terms that best describe the mental-emotional state of the photographed actor.	Reading the Mind in the Eyes (RME): Baron-Cohen et al. (1997).	Affective	Non-verbal
Static FB protocols. Object displacement tasks (e.g., Sally-Anne or Maxi task) - inferring mental states of naïve agents precluded from privileged knowledge.	Protagonist places an object in location X and then leaves the scene. In their absence, the object is moved from location X to location Y. The protagonist returns to the scene and then the participant is asked the critical question: "Where will he/she look for the object?" To resolve the task, the participant must represent the protagonist's false belief (that the object is in location X) and inhibit their own true belief (object is in location Y).	Wimmer and Perner (1983).	Cognitive	Non-verbal
Inference of mental states of characters in dynamic, ecological social scenarios	Short films depicting social scenarios. Film is periodically paused and participants are asked questions related to characters' mental states. Purported to possess greater ecological properties compared to static paradigms.	Movie for Assessment of Social Cognition (MASC): Dziobek et al. (2006).	Cognitive / affective	Non-verbal
Inference of characters' mental states in ToM	Participants are presented with a sequence of comic-book-style stills depicting a social interaction. Participants are asked to infer/attribute first- and second-order FBs in addition to answering	ToM Picture Story Task: Brüne (2003).	Cognitive	Non-verbal

storyboard or comic-strip tasks.	non-social control questions. In some variants, participants are required to 'finish' the story by selecting the most appropriate ending to the comic-strip sequence based on inference of characters' mental states.			
Imputation and inference of mental states in non-human, animated shape forms.	Participants shown a video of two animated triangles moving in a manner that implies intentions (e.g., pretending, deceiving, etc.). Participants asked to describe the intentionality of the presented triangles, where 'appropriateness' (given the implied social context) is the outcome measure.	Geometric Triangles Animation: Castelli (2002).	Cognitive	Non-verbal

Note: This is not intended to be an exhaustive list of all published ToM protocols, rather it is a summary of a selection of the most popular experimental ToM paradigms. Due to task heterogeneity (in modality, domain, administration [online, 'in-person'], experimenter involvement, and outcome measure), it is not possible to neatly taxonimise all variants and iterations of the above-mentioned tasks into single, clearly defined categories. For a breakdown of paradigms according to ToM 'mechanisms', see Byom & Mutlu (2013).

Children below age 4-5, who have as yet immature neural systems to support some EFs, typically 'fail' false belief tasks. It has been suggested that this 'failure' reflects processing demands rather than a specific inability to represent and mentalize about false beliefs (Scott & Baillargeon, 2017). Similarly, in advanced age, there is an established literature detailing age-related vulnerabilities in cognitive function, particularly processes that are traditionally thought to rely on the prefrontal cortex (West, 1996; 2000). Models such as the *disconnected brain theory* (e.g., Fjell et al., 2017) or the *frontal lobe hypothesis of ageing* (e.g., Cabeza & Dennis, 2012; West, 2000) generally posit a framework through which cognitive ageing can be understood and indexed. Accounts of age-related structural and functional deficits in the brain, such as changes in connectivity between regions that are important for particular executive abilities – i.e., the inferior frontal gyrus with respect to inhibitory control (Banich and Depue, 2015) – have advanced our understanding of the changes in the respective neural mechanisms that underlie particular aspects of cognitive ageing. EF, commonly conceptualised as inhibitory control, attention switching or set-shifting, and working memory updating (Miyake et al., 2000), have been reliably shown to decline with age (for reviews, see Harada, Love, & Triebel, 2013; Salthouse, 2010; 2019). For example, in a recent longitudinal investigation of 781 older adults, Caballero et al. (2020) used eight different neuropsychological indicators to map changes in the above-mentioned EFs. The authors reported significant longitudinal decline with increasing age, where the decline with age was more prominent in older-old participants. Interestingly, and more of which later, after age, the variability in the downward trajectory of EF was related to level of education.

More recently, Maldonado et al. (2020) meta-analysed age differences in EF on multiple levels, including task level (e.g., Stroop, Flanker, Trail Making Test, etc.) and a domain and global EF level (inhibition, updating, shifting, and processing speed). In their analysis of 16 different EF tasks, a robust age difference emerged (Hedge's $g = 1.29$); though it is worth noting that three task-types yielded a non-significant age effect: Go/No-Go, Task Switching, and Letter Comparison tasks. More broadly, strong effects were also observed on the domain level, ranging from .8 in Updating to 1.64 in Inhibition. While there is variability in the magnitude of an age effect depending task type, the literature generally points toward an age-related decay of EFs in later life.

With respect to SCF, there is noticeable disparity across ToM tasks (Quesque and Rossetti, 2020). Indeed, there are numerous methods with which to assess ToM; however, some protocols may be more burdensome than others due excessive reliance on supporting language and EF systems. As we have seen above, healthy ageing brings about wide-ranging declines across multiple facets of EF. Taken together, as typical ToM protocols involve EF demands to varying degrees, and as EFs have been shown to reliably wane with age, conventional ToM protocols may inadvertently inflate age-related effects through non-social (i.e., EF-related), incidental task demands that disproportionately disadvantage older adults. Likewise, the nature of ToM tasks, with respect to how stimuli are constructed and presented, and their subsequent differential cognitive demands, may also in part drive heterogeneity in age effects.

ToM tasks can be broken down into verbal and non-verbal modalities, where verbal tasks present stimuli primarily in textual form, comprising stories or vignettes (e.g., Happé et al., 1998; Saxe and Kanwisher, 2003). Non-verbal tasks, on the other hand, are primarily picture-based and depict ToM scenarios through static comic-strips or images (e.g., Mind in the Eyes task, Baron-Cohen et al., 1997) or movies/animations (e.g., Movies for Assessment of Social Cognition, MASC; Dziobek et al., 2006). Given the differential prominence of language content between these two task modalities, it is plausible that individual differences in language ability may, at least in part, explain variability in ToM performance. This is an important consideration when examining an ageing-ToM effect as particular faculties of language have also been shown to change with age. In terms of language production, OA typically produce syntactically simpler speech in comparison to their younger counterparts and have more frequent (and longer) empty pauses (Kemper & Sumner, 2001). Further, OA typically show more difficulty with word finding, and have been shown to produce more 'tip of the tongue' states (a language *failure*) in comparison to YA (Juncos-Rabadán et al., 2010; Segaert et al., 2018). With respect to language comprehension, the evidence of an age difference is less clear. Studies have shown that core aspects of comprehension are preserved and even improved in advanced age, however, significant variability exists depending on the method by which performance is assessed (for reviews, see Schneider, Pichora-Fuller, & Daneman, 2010; Thornton & Light, 2006). Most pertinent to text-based ToM paradigms is perhaps OA' reading and word recognition abilities. Behavioural reading speed studies have shown that OA exhibit up to 30% slower reading speeds relative to their younger counterparts (Liu, Patel, & Kwon, 2017; Rodriguez-Aranda, 2003). Similarly, eye-tracking studies have also found that OA show slower lexical recognition across a range of textual stimuli compared to YA (e.g., see Kliegl et al., 2004). Finally, the age effect in reading and word recognition speed seems also to vary as a function of *word frequency*: how common and/or well-known a word is within a language. Overall, OA are slower and find more difficulty with accessing less well-known words relative to YA (DeDe & Flax, 2016).

The differences in language processing between young and older participants is an important consideration when examining an overall ToM age effect. Verbal ToM tasks typically present text-based stimuli at relatively (and un-naturalistically) fast speeds.

The content of the stimuli itself also qualitatively varies in lexical difficulty. Consider, for example, Saxe and Kanwisher's (2003) original ToM localizer task. This task comprises ToM vignettes that depict either mental content that requires mentalizing about the knowledge states of characters (i.e., FB trials), or vignettes depicting outdated physical content that can be resolved based on mechanical or physical inference (i.e., True Belief trials; Saxe & Kanwisher, 2003). While the ToM localizer has been used extensively in the ToM literature, the authors of the task, however, did not account for word frequency or syllable count when constructing their task stimuli. For example, the relatively infrequent and somewhat specialist word, *petroglyphs*, appears in one vignette in the original Saxe and Kanwisher stimuli set. Indeed, the majority of vignette-based paradigms in the literature do not account for lexical difficulty (word frequency, syllable count, sentence length, etc.). This highlights a broader point about the lack of carefully controlled stimuli and wider experimental design in classic ToM paradigms. It is plausible, then, given the evidence presented with respect to age-related modification of language abilities, and the variability in lexical difficulty across ToM (especially, *verbal* ToM) tasks, that an age effect in ToM emanating from these protocols may in part reflect non-social language and EF changes.

The heterogeneity in ToM age effects may also reflect sampling and methodological design. First, in terms of investigating an age effect, what constitutes 'advanced age' or 'old' appears to vary from study to study. For example, in a recent study by Calso et al. (2020) where the authors tested young vs. older adults on a picture-based story paradigm that depicted social scenarios of deception and cooperation, the mean age of the older adult cohort was over 84. In contrast, Zhang et al. (2018) tested young and older adults on an animated shape form paradigm, where participants were asked to describe the intentionality of the presented triangles. The older sample in

Zhang et al. was aged, on average, around 65. Further, to address the issue of within-group age heterogeneity in older adults, studies have commonly subdivided older participants into young-old and old-old groups. However, even with these smaller, constrained age groupings, there is a slight (but meaningful) inconsistency in the literature as to what constitutes young-old and old-old. For example, Maylor et al. (2002) and Lecce et al. (2019) used a threshold of 75+ years to demark old-old, whereas Calso et al. (2020) used a higher cut-off of 80+. Overall, the variation in age in older participants groups is an important consideration when assessing the effect of age on ToM. As we have discussed previously, EF – abilities which typically drive aspects of ToM performance – decline with age, and this age-related modification is most severe in the older-old groups (Salthouse, 2010; 2019). Therefore, due to these sampling inconsistencies, it is difficult to make reasonable comparisons between studies to discern a *true* effect of age, as the effect of age seems to be confounded with EF, language, and sampling inconsistencies.

Aside from an age consideration, there is also heterogeneity in the ToM literature with respect to OA' education. For example, in the previously-discussed Calso et al. (2019), the average education of the OA sample was 9.8 years, whereas Lie et al. (2013) reported an average age of almost 16 years. The disparity in education is an important factor when assessing age-related changes in SCF (and indeed, cognitive ageing more generally) as education has been shown to be protective against some forms of cognitive decline associated with advanced age. For instance, in a non-parametric quantitative review of 18 different studies comprising over 47 000 individual participants, Valenzuela and Sachdev (2006) examined the relationship between education and longitudinal cognitive decline. The authors reported a significant and large effect ($\varphi = 2.63$) of education on cognitive decline, where lower levels of education were associated with greater magnitudes of deterioration. More recently, Opdebeeck, Martyr, and Clare (2019) conducted a meta-analysis to examine the association between cognitive reserve (i.e., cognitive flexibility) and education, occupation, and engagement in cognitively-stimulating activities, in healthy older adults. The authors reported a modest, significant association between level of education and measures of working memory, executive function, visuospatial ability, and language.

In terms of the link between education and intelligence in older age, there is divergence with respect to the domains of intelligence. There is general agreement that fluid intelligence (cognitive adaptation, novel problem-solving, etc.) declines with age, and that this decline is somewhat moderated by years of education (see meta-analysis by Ritchie & Tucker-Drob, 2018). Crystallised intelligence, conversely, seems to be preserved and even improved with age (Horn & Cattell, 1967; Kaufman & Lichtenberger, 2006; Salthouse, 2004). Evidence has shown that, for example, brain regions that are involved in mediating the storage of fact and object-based information – e.g., the parahippocampal cortex – show resistance to typical age-related structural changes (Aminoff, Kveraga, & Bar, 2013). Kaufman, Kaufman, and Johnson (2009) examined the association between educational attainment and fluid and crytsallised intelligence in over 1000 adult participants, ranging in age 22 to 90. Participants' education was categorised as 8th-11th Grade, High School Graduate, Some College, and 4-year Degree. After covarying out the effect of age, Kaufman and colleagues reported significant associations between education level and both fluid and crystallised intelligence. More specifically, adjusted for age, participants who had obtained a *4-Year Degree* exhibited significantly greater scores on measures of fluid and crystallised intelligence relative to their lower-educated counterparts. In addition to measures of intelligence, the authors also administered measures of academic ability (math, writing, and reading) and then also related these to educational attainment. Interestingly, and related to a previous discussion about language abilities in older age and a possible confound in verbal ToM tasks, the highest academic abilities (of which reading is pertinent here) were found in those participants with the highest level of education. Intelligence is an important factor when considering ageing ToM. Studies have shown that intelligence can, at least in part, predict SCF. For example, Ibanez et al. (2013) reported that intelligence (in particular, fluid intelligence) was found to partially explain individual differences in ToM performance. It should be noted, however, Ibanez and colleagues tested only affective ToM (via Baron-Cohen et al.'s [2001] RMET).

In classic studies of ageing ToM, however, education is seldom carefully considered. Further, as we have seen above, wider demographic and cognitive variables, such as age (of OA), language abilities, and intelligence, that have been shown to either directly or indirectly influence SCF, are also largely overlooked when comparing young vs old ToM. Thus, an uncorrected like-for-like comparison of existing ToM studies, where there is substantial variability in the above-mentioned factors, may potentially inflate an effect of age. It is important, therefore, to explore the specificity of an age effect across and within studies with respect to a wide range of factors beyond those ordinarily included in studies of ageing ToM.

Purpose and scope of current review

The purpose of the current review was to address three objectives. **1)** First, we aimed to build upon and extend the reviews of Henry et al. (2013), Moran (2013), and Moran, Jolly, and Mitchell (2012) to interrogate the literature-base since the publication of these papers. While there has been a marked increase in publications investigating ToM in ageing since 2013 (see Figure 2.1), to our knowledge, there has been no meta-analysis of young vs. old behavioural, cognitive ToM since the release of the above-mentioned reviews. It is our aim, therefore, to chart the extant ToM-ageing data and provide a contemporary analysis and highlight the nature and magnitude of the much-reported age-related decline in ToM.

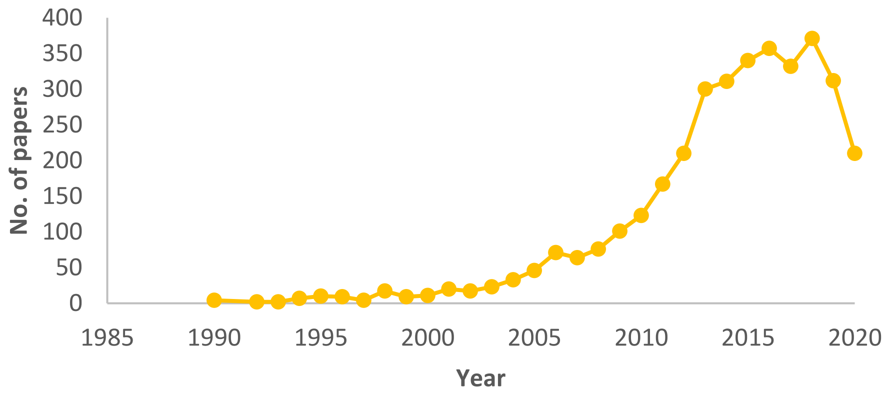


Figure 2.1. Growing trend of ToM-ageing publications. Data represents number of hits in response to the search query ((“theory of mind” or “social cognition”) and “ag?ing”) on PubMed (December 2020).

2) Second, we set out to aggregate and quantify the magnitude of a potential age effect within only *cognitive* ToM tasks. Previous meta-analyses have included both cognitive and affective domains when computing an overall age effect. In contrast to Henry et al. (2013), we only included experiments where the critical aspect of the task involved reasoning or mentalizing about knowledge, belief, or intentional states (i.e., cognitive ToM). Indeed, the literature often reports differential age effects for cognitive and affective ToM, where, for example, Bottiroli et al. (2016) reported selective age-related deficits only in cognitive – but not affective – ToM. At the time of writing, there is no such analysis of cognitive ToM in ageing; as such, we aimed to chart the magnitude of an age-related difference in measures of cognitive ToM. **3)** The final objective of this study was to examine the specificity of any age effect within specialised, theoretically-driven sub-group analyses. We aimed to investigate the potential effect of age as a function of task domain, modality, and participant characteristics, namely, OA age and education. These analyses were performed to address the insufficient clarity in the current literature about the specificity of an age effect. We expected that if there was indeed an effect of age, there would be heterogeneity in its magnitude according to task and participant characteristics. This was based on work suggesting that there is large variation in the neuropsychological demands across different tasks, where some tasks tap domain general, non-social resources more than others, for example, some tasks may require language comprehension and reading abilities for successful performance (e.g., Happé et al.'s *Strange Stories*), other tasks typically may mask non-social demands within social manipulations – e.g., attentional resources necessary to keep track of contents of revealed locations in typical Sally-Ann-style ToM paradigms. It was our aim, therefore, to disentangle and highlight the separate influences of task demands and participant characteristics in a set of specialised series of sub-group analyses, through theoretically-driven task taxonomies and participant (demographic) variables.

Method

Where applicable, the current analysis was conducted in adherence with the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA; Moher et al., 2009).

Literature Search

A review of the literature was performed in May 2020 by searching the *Web of Science* (Thomson Reuters), *Psych INFO* (American Psychological Association), and *PubMed* (United States National Library of Medicine) databases. The reference lists of included studies were also scrutinised to identify additional relevant papers. Grey literature and unpublished theses were precluded from the study.

The search parameters were based on earlier reviews by Henry et al. (2013) and Shurz et al. (2014). In the current review, the key terms were centred on two core areas: 1) theory of mind (e.g., theory of mind, mentalizing, false belief, belief reasoning, mindreading, perspective-taking, and 2) ageing (e.g., ageing, older adults, elderly, and old). The search strategy, including the use of wildcards, was as follows:

> ('theory of mind' OR 'mentaliz*' OR 'false belief' OR 'belief reasoning' OR 'mind?read*' OR 'perspective-taking') AND ('ag?ing', 'older adult*', 'elderly', 'old*'.)

No age limiter was applied to the search.

Inclusion Criteria

Inclusion and exclusion criteria were determined prior to performing and examining the results of the literature search. Studies were included based on the following criteria: 1) Comparison of healthy young vs. older adults, with 'healthy' defined as the reported absence of any neuropsychological disorder. 2) Community-dwelling older groups (i.e., not institutionalised). 3) Behavioural measure of ToM ability must have been based on *cognitive* (not *affective*) ToM. 4) At least one form of ToM behavioural measure reported with a young vs. old comparison. 5) 'Older' defined as adults aged >60. 6) Sufficiently precise (*M* and *SD* for cognitive ToM measure, per age group) reporting of statistics as to enable calculation of effect sizes. 7) Papers must have been written in English.

Exclusion Criteria

Literature was rejected based on the following exclusion criteria: review papers, conference proceedings/abstracts, non-peer reviewed literature, grey literature, and papers published in a language other than English.

Screening and Data Extraction

Titles, abstracts, and full-texts were screened by the author and an independent reviewer (Dwayne May, Aston University), and disagreements were resolved in meetings to reach consensus. The following data were extracted from the included studies for data synthesis: author(s) and year of publication, country, number of participants (total and per group), descriptives (*M*, *SD*) of age per group, sex ratio (if reported), education (years) per group, ToM task type, and ToM performance (*M* & *SD* of accuracy of ToM task).

As with Henry et al. (2013), only accuracy (not response-time measures) were permitted for analysis. For studies that employed a 'lifespan' approach that fractionated age into smaller, more granular groups, only the youngest and oldest groups contributed to the analysis (provided that the groups met the Inclusion Criteria outlined above). Studies that did not group age into discrete young and old categories – i.e., studies that track the trajectory of ToM ability with age in a continuous or longitudinal manner – were not permitted. To be consistent with the wider ageing literature, healthy ageing was defined as individuals aged 60 and over with no current diagnosis of severe or life-limiting morbidity. Younger adults were defined as healthy individuals aged 18+ (and less than 60); though, in reality, the majority of studies used university students as their younger sample, which naturally skews the average age of the younger group toward the early 20s. In the context of multi-study papers, if only some experiments met the inclusion criteria, these were included and those experiments that that did not fit the criteria were excluded. Likewise, where studies used ToM tasks that comprised both a cognitive and affective component (e.g., the Yoni task, as in Fischer, O'Rourke, & Thornton, 2017), only the cognitive ToM data contributed to the analysis. Some studies reported both unadjusted and adjusted ToM performance data (i.e., with complementary measures of executive function and/or intelligence; for example Slessor, Phillips, & Bull, 2007), whereas others reported only unadjusted scores (e.g., Bernstein, Thornton, & Sommerville, 2011). For consistency, only the unadjusted data were permitted for analysis. Self-reported, subjective measures of social cognitive ability were not permitted (e.g., the 'Perspective-Taking' component of the Interpersonal Reactivity Index; Davis 1980). Finally, for studies that reported more than one outcome for the same cognitive ToM task (e.g., Maylor et al., 2002), data were pooled to create a unitary effect size estimate. For studies that reported more than one outcome from different cognitive ToM tasks, two approaches were taken: 1) for the overall meta-analytic effect, data were pooled across tasks to create a single effect size (as above); and 2) for the sub-groups analysis, task data were grouped according to type and if there was non-individuality or statistical dependence in the data, e.g., where multiple (different) cognitive ToM experiments are reported from the same cohort of participants, a conservative approach was taken where the ToM experiment which yielded the smallest effect was included. These data processing decisions align with previous published cognitive ageing meta-analyses (e.g., Henry et al., 2013; Rey-Mermet & Gade, 2018).

Statistical Analysis

Effect sizes were computed based on standardised mean difference (SMD) between the two age groups. SMD is statistically identical to Cohen's *d*, and thus effect sizes are reported as *d* henceforth. Interpretation of effect sizes was based on the conventional thresholds of Cohen's d: small (d = 0.20), moderate (d = 0.50), and large (d = 0.80). Effect sizes were calculated using the conventional formula:

$$\frac{\text{Mean 1} - \text{Mean 2}}{\text{Pooled SD ([SD1 + SD2]/2)}}$$

The significance and extent of heterogeneity was calculated using Cochran's Q statistic and I^2, respectively. Cochran's Q, distributed as a chi-square statistic, is calculated as the weighted total of squared differences between each study effect and the pooled effect across studies (Hoaglin, 2016). Relatedly, the I^2 is a measure of the variability (in percentage terms) in the estimated effect that is based on heterogeneity as opposed to sampling error or chance. A rough but useful guide in interpreting I^2 is as follows: 0% to 40%: might not be important; 30% to 60%: may represent moderate heterogeneity; 50% to 90%: may represent substantial heterogeneity; and 75% to 100%: considerable heterogeneity (see Higgins, 2003).

In a meta-analysis, when computing an overall 'true' effect size, there are two fundamental models that can be adopted: the fixed-effect model and the random effects model (for a detailed account of fixed vs. random effects models in meta-analyses, see Hunter & Schmidt, 2000). A central process of a meta-analysis is the computation of the *combined effect* (or the overall, mean effect size). As there are differences in the precision of the estimated means across studies, a simple unweighted overall mean of effect sizes is likely to inflate or diminish any *true* effect. To account for this, means are typically weighted according to level of precision (sampling error), where studies with larger Ns typically being assigned greater weight. The designation of these weights is where the two previously mentioned approaches (fixed vs. random) essentially differ. With a fixed-effect model, there is a core assumption that there is one universal *true* effect that underlies all studies included in the analysis. With a random effects model, however, the aim is to estimate the mean of a distribution of 'true' effects. Studies with larger Ns may return more precise estimates than smaller studies; and each study is effectively estimating a dissimilar effect size where each of these effects serve as a sample from the population under investigation. Therefore, compared with the fixed effect model, where there is the assumption that all studies compute qualitatively the same effect, the designation of weights under random effects are better adjusted. Larger studies are less likely to have an overbearing influence in the analysis and smaller studies are less likely to be underestimated. For these reasons, a random effects model was used in the current analysis. This is in line with previous ageing ToM meta-analyses (e.g., Henry et al., 2013).

Publication Bias

Meta-analyses aggregate data from a large corpus of studies and thus benefit from higher statistical power and more robust point estimates compared to individual experiments. However, within the pipeline of conducting a meta-analysis, the researcher makes various judgements which may ultimately affect the outcome of the overall analysis. These decisions include (but are not limited to) deciding on and implementing the literature search strategy, devising inclusion and exclusion criteria, resolving incomplete data, and determining the parameters of the statistical analyses. Another important consideration in undertaking a meta-analysis is the issue of publication bias, where the outcome of a study may influence the decision whether or not it is published (Begg, 1994). This bias can take one of many forms: deliberate non-publication of non-significant results, partial (selective) reporting or 'cherry picking' of results, manipulation or 'massaging' of data to transform non-significant results into nominally significant outcomes (Ioannidis, 2017). As a result of these practices, the body of literature available when conducting a summary of the evidence is potentially not a valid reflection of all work undertaken in the given discipline. Due to the 'file drawer problem', where non-significant findings are less likely to be written up for submission or go on to be published, there may be a biased distribution of effect sizes where the magnitude of the results of published studies are exaggerated. These are important considerations as the selective reporting of *positive* (statistically significant, in the expected direction) results pose a potentially serious threat to the validity of a meta-analysis. The detection of potential publication bias within the context of a meta-analysis is therefore crucial.

One such method of detecting publication bias is through the visual examination of funnel plots. The funnel plot is essentially a scatterplot of effect size against a measure of sampling precision (typically indexed as standard error). When there is little-to-no publication bias, the plotted effects form a symmetric inverted funnel shape. Asymmetry, or a shape that is not consistent with the conventional funnel pattern, is interpreted as indication of publication bias. The inspection of funnel plots is visual and therefore subjective in nature. While more sophisticated and objective statistical methods exist for the detection of publication bias (e.g., Egger's regression test, see Egger et al., 1997), a visual examination of funnel plots is sufficient for non-complex meta-analyses in the first instance. Accordingly, in the current study, a funnel plot was used to detect potential publication bias.

Analysis Software

The desktop version of the open-source RevMan 5 (version 5.4.1) meta-analysis package was used for the current analysis. RevMan is an analysis tool developed and maintained by the Cochrane Collaboration for the purposes of formulating systematic reviews and meta-analyses. In addition, Metafor – a meta-analysis package for *R* (run in RStudio v. 3.5.0) – was used for data visualisation. Coding and manipulation of raw data was performed in Microsoft Excel.

Results

Study Selection

The initial searches returned 426 articles, of which 178 full-text articles were assessed for eligibility; see Figure 2.2 for PRISMA flow diagram of identified articles. Overall, 28 articles that comprised 35 datasets met the inclusion criteria, with a total of 1151 (53%) YA and 1033 (47%) OA included in the analysis. Within the YA cohort, there were 453 (41%) male and 643 (59%) female participants, and in the OA cohort, 383 (39%) males and 600 females (61%). Two studies (Cavallini et al., 2013 & Duval et al., 2011) did not report sex data. The YA cohort had a mean reported age of 22.34 (SD = 2.59, range 18 – 30) and OA had a mean age of 73.46 (SD = 4.95, range 60 – 84). In terms of education, the YA cohort reported a mean of 15.02 years (SD = 1.36) and OA reported 13.34 years (SD = 2.10). Education data were unobtainable from nine studies either due to omission (where study authors were contacted and there was no response) or education was reported as a categorical variable (e.g., “Undergraduate Bachelor’s Degree”) which prevented the computation of descriptive statistics. These studies were Franco et al. (2013), Grainger et al. (2020), Johansson Nolaker et al. (2018), Moran et al. (2012), Phillips et al. (2011), Rakoczy et al. (2012), Reiter et al. (2017), Sullivan & Ruffman (2004), and Zhang et al. (2018).

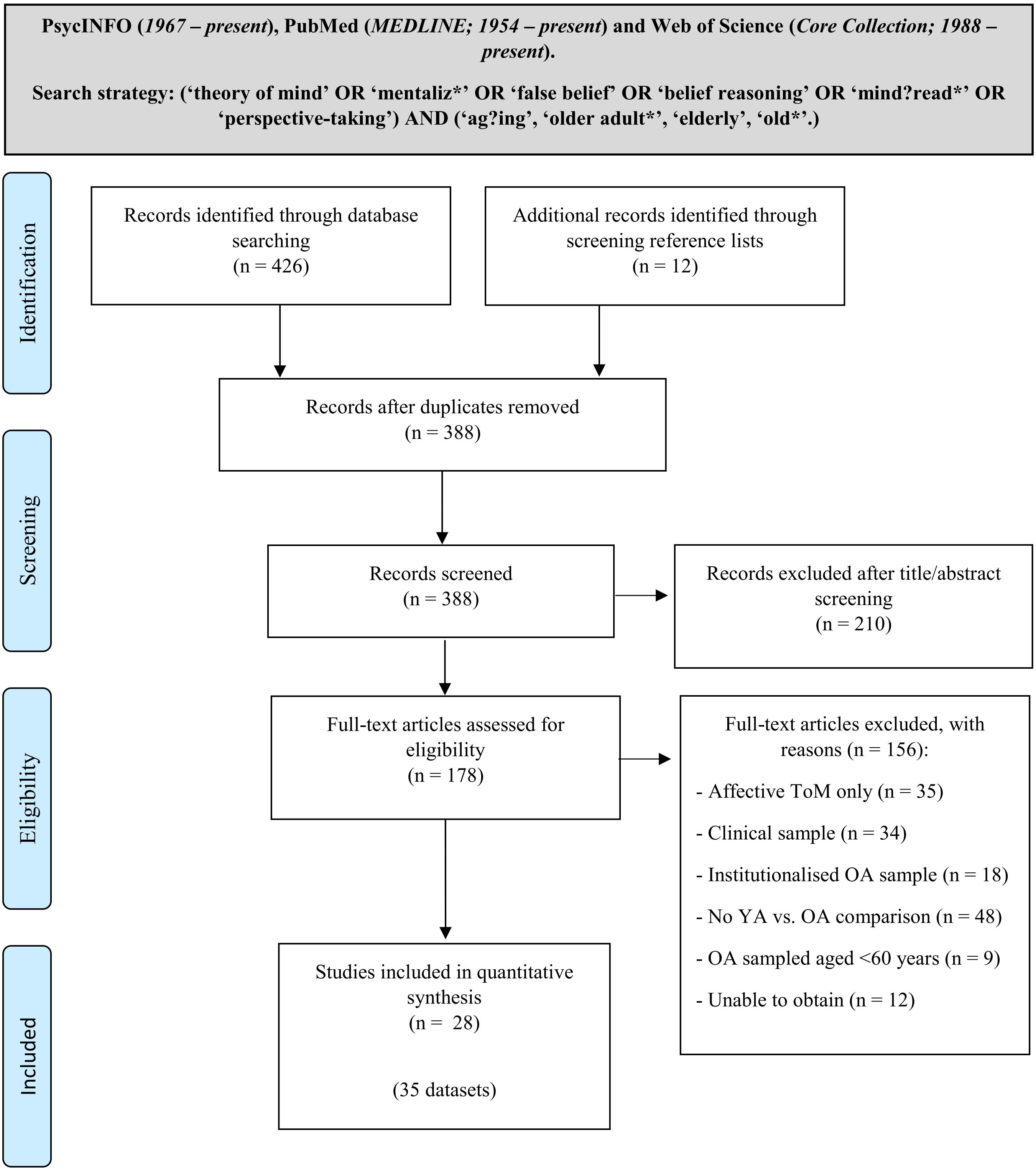


Figure 2.2. Search strategy and PRISMA flow diagram of identification and screening of articles. OA = older adults. YA = younger adults.

For each included study, author(s), year of publication, country, total N, N per group, age M/SD, sex ratio, task information (task name, modality, paradigm), and effect size (standardised mean difference, expressed as *d* and its associated 95% CI), are presented in Table 2.2. Studies that comprised more than one ToM experiment/outcome are marked with an asterisk (*).

Table 2.2.
Included studies and their respective participant and task characteristics.

		Participant characteristics						ToM task information			Effect size (SMD)	
		n		*M age*		Sex (M:F)						
Study	Country	YA	OA	YA	OA	YA	OA	Task & task origin	Modality	Paradigm note	*d*	95% CI
Baksh et al. (2020)	UK	30	31	22.57	72.29	12:18	16:15	ESCoT; Baksh et al. (2018)	Non-verbal; animation (film)	Computerised animation of social interactions. Comprises both cognitive and affective ToM; only cognitive ToM included in current analysis.	.97	.43 – 1.50
Bernstein et al. (2011)	Canada	37	37	19.20	67.6	9:28	9:28	Sandbox Task	Non-verbal; static cartoon / comic strips	Based on classic Sally-Ann paradigm; participants physically point to 'correct' location. Spatial distance between participant response and correct response taken as DV.	.98	.49 – 1.46
Bottiroli et al. (2016)	Italy	20	20	22.75	75.50	9:11	6:14	Faux Pas; Stone & Baron-Cohen, (1998)	Verbal; stories/vignettes	Detection of faux pas in social scenarios. Comprises both cognitive and affective components – only cognitive ToM data used in current analysis.	1.32	.63 – 2.01
Calso et al. (2019)*	France	30	20	25.60	83.30	15:15	5:15	MPS-TOMQ	Non-verbal; static cartoon / comic strips	Picture stories involving deception and social reciprocity. Participants required to arrange pictures in logical order, and then respond (detect) deception, cheating, reality, and first-/second-order belief.	1.32	.69 – 1.77
								TOM-15; Desgranges et al. (2012)	Non-verbal; static cartoon / comic strips	First- & second-order FB task comprising picture stories; participants required to choose correct response from two possible options. Control condition tests non-mental aspects of stories.	1.72	1.05 – 2.38
Calso et al. (2020)	France	35	30	25.43	84.37	17:18	5:25	ToM Picture Story; Brüne (2003)	Non-verbal; static cartoon / comic strips	Picture stories depicting social scenarios of deception and cooperation; first- and second-order FB understanding also tested.	1.92	1.32 – 2.51

(Table continues)

Cont'd		Participant characteristics						ToM task information			Effect size (SMD)	
		n		*M age*		Sex (M:F)						
Study	Country	YA	OA	YA	OA	YA	OA	Task & task origin	Modality	Paradigm note	*d*	95% CI
Calso et al. (2020)	France	35	30	25.43	84.37	17:18	5:25	ToM Picture Story; Brüne (2003)	Non-verbal; static cartoon / comic strips	Picture stories depicting social scenarios of deception and cooperation; first- and second-order FB understanding also tested.	1.92	1.32 – 2.51
Cavallini et al. (2013)	Italy	30	29	23.63	74.99	NR	NR	Strange Stories; Happé (1994)	Verbal; stories/vignettes	Detection of protagonists' intention, irrespective of reality. Stories contain double-bluff, lie, joke, sarcasm, and pretence.	2.67	1.96 – 3.38
Duval et al. (2011)*	France	25	25	23.80	70.14	NR	NR	Attribution of Intention Task; Brunet (2000)	Non-verbal; static cartoon / comic strips	Comic strips depicted social scenes. Participants selected most logical conclusion for each scenario. Two conditions: character intention (ToM) or control scenarios (properties of objects).	.02	-.54 – .57
								FB object-displacement (e.g., 'Sally-Ann'); Wimmer & Perner (1983)	Non-verbal; static cartoon / comic strips	Comic strips followed by test questions, comprising true and false belief scenarios. First- and second-order FB. Requires inference of naïve protagonists' (outdated) knowledge/mental state.	.29	-.27 – .85
El Haj et al. (2016)	France	40	36	23.13	69.53	18:22	16:20	FB object-displacement; Desgranges et al. (2012)	Non-verbal; static cartoon / comic strips	Variation of classic categorical object-location FB task, based on Sally-Ann paradigm (Wimmer & Perner, 1983).	.84	.37 – 1.31

Cont'd		Participant characteristics						ToM task information			Effect size (SMD)	
		n		*M age*		Sex (M:F)						
Study	Country	YA	OA	YA	OA	YA	OA	Task & task origin	Modality	Note	*d*	95% CI
Franco et al. (2013)	Colombia	83	89	19.10	69.40	51:32	39:53	Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	.05	-.25 – .35
Fischer et al. (2017)*	Canada	86	85	19.80	71.4	23:63	26:59	Yoni Task	Non-verbal; static cartoon / comic strips	Depiction of a fictional character – "Yoni" – that is presented with four images from which participants chose most appropriate mental or affective state.	.74	.43 – 1.05
								Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	1.08	.76 – 1.41
German & Hehman (2006)	USA	27	20	20.00	78.00	9:18	1:19	ToM Belief-Desire Stories	Verbal; stories/vignettes	Participants exposed to stories involving belief-desire reasoning problems, and control stories with no mental state content.	.42	-.16 – 1.01
Grainger et al. (2018)	Australia	51	50	20.45	71.67	19:32	24:26	FB Movies; Schneider et al. (2013)	Non-verbal; movies	Eye-tracking study using traditional Sally-Ann-style (object-location), categorical scenarios in film form. Only behavioural responses included in current analysis.	.03	-.36 – .42
Grainger et al. (2019)	Australia	48	50	20.67	75.38	15:33	17:33	TASIT; McDonald et al. (2011)	Non-verbal; movies	Videotaped vignettes of social interactions; testing the detection of intention, attitudes, and meaning.	1.75	1.28 – 2.22

Cont'd		Participant characteristics						ToM task information			Effect size (SMD)	
		n		*M age*		Sex (M:F)						
Study	Country	YA	OA	YA	OA	YA	OA	Task & task origin	Modality	Note	*d*	95% CI
Grainger et al. (2020)*	Australia	40	40	19.00	70.73	20:20	20:20	VAMA – Virtual Assessment of Mental Attribution	Non-verbal; interactive, immersive	Computerised, first-person virtual shopping task with virtual friends. Participants perform shopping 'errands', and answer questions about interactions involving their virtual friends.	.47	.02 - .91
								Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	.10	-.34 – .54
Happé et al. (1998)	UK/USA	67	19	21.75	73.00	33: 34	9:10	Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	-1.12	-1.66 – -.58
Hughes et al. (2019)	USA	40	35	21.58	71.66	15:25	13:22	ToM Localizer; Saxe & Kaniwisher (2003)	Verbal; stories/vignettes	Task has two conditions: False belief and False story, requiring participants to mentalize or make mechanical/physical inferences, respectively.	.48	.02 – .94
Lecce et al. (2019)*	Italy	30	31	21.97	78.87	9:21	10:21	MASC – Movies for Assessment of Social Cognition; Dziobek et al. (2006).	Non-verbal; movies	Short, ecological films depicting social scenarios. The film is periodically paused and participants are asked questions related to characters' mental states.	2.46	1.79 – 3.14
								Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	1.03	.50 – 1.57

Cont'd		Participant characteristics						ToM task information			Effect size (SMD)	
		n		*M age*		Sex (M:F)						
Study	Country	YA	OA	YA	OA	YA	OA	Task & task origin	Modality	Note	*d*	95% CI
Li et al. (2013)*	China	28	28	20.46	76.29	13:15	15:13	False Belief Vignettes; Based on Wimmer & Perner (1983)	Verbal; stories/vignettes	Short stories based on Wimmer & Perner's (1983) object-location paradigm. Mix of ToM and control (physical causation) stories.	.01	-.52 - .52
								Faux Pas Test; Stone & Baron-Cohen (1998)	Verbal; stories/vignettes	Detection of faux pas in social scenarios. Comprises both cognitive and affective components – only cognitive ToM data used in current analysis.	.07	-.45 - .60
Maylor et al. (2002)	UK	55	55	20.10	80.80	29:26	19:36	Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	2.01	1.38 – 2.63
Moran et al. (2012)	USA	29	13	23.00	71.80	10:19	7:6	ToM Localizer; Saxe & Kaniwisher (2003)	Verbal; FB vignettes	ToM Localizer task as described previously.	.66	-.01 – 1.33
Nolaker et al. (2018)	UK	20	19	21.00	70.00	3:17	10:9	Strange Stories Film Task; Murray et al. (2017)	Non-verbal; movies	Ecological ToM films based on Happé et al.'s (1994) Strange Stories. ToM films involved telling a white lie, double bluff, & persuasion.	.91	.25 – 1.58
Phillips et al. (2011)	UK	52	36	25.81	73.67	25:27	15:21	ToM Videos Task	Non-verbal; movies	Cartoonised films of a ToM object-displacement task with three possible target locations, testing first-order FB.	1.04	.58 – 1.49

Cont'd		Participant characteristics						ToM task information			Effect size (SMD)	
		n		*M age*		Sex (M:F)						
Study	Country	YA	OA	YA	OA	YA	OA	Task & task origin	Modality	Note	*d*	95% CI
Rakoczy et al. (2012)	Germany	27	20	22.67	73.3	13:14	11:9	Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	.80	.19 – 1.40
Rakoczy et al. (2018)	Germany	40	40	24.35	68.43	15:25	17:23	Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	.61	.61 – 1.06
Reiter et al. (2017)	Germany	55	52	24.29	72.08	24:31	22:30	EmpaToM; Kanske et al. (2015)	Non-verbal; movies	Comprises cognitive & affective ToM, and empathy/compassion. 15-s video clip of actor narrating an autobiographical experience. Participants answer questions pertaining to first-order FB.	1.21	.79 – 1.62
Slessor et al. (2007)	UK	40	40	20.08	66.95	12:28	15:25	Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	.21	-.23 – .65
Sullivan et al. (2004)	UK	24	24	30.00	73.00	13:11	8:16	Strange Stories; Happé (1994)	Verbal; stories/vignettes	Strange Stories task as previously described.	1.36	.73 – 1.99
Zhang et al. (2018)*	China	87	84	25.62	65.47	37:50	34:50	Geometric Shapes Animation; Castelli (2002)	Non-verbal; animation (film)	Participants shown a video of two animated triangles moving in a manner that implies intentions (e.g., pretending, deceiving, etc.). Participants asked to describe the intentionality of the presented triangles.	.92	.60 – 1.23

Cont'd		Participant characteristics						ToM task information			Effect size (SMD)	
		n		*M age*		Sex (M:F)						
Study	Country	YA	OA	YA	OA	YA	OA	Task & task origin	Modality	Note	*d*	95% CI
Cont'd - Zhang et al. (2018)*								Faux Pas Test; Stone & Baron-Cohen (1998)	Verbal; stories/vignettes	Faux pas test as previously described.	.23	-.07 – .53

Note: *More than one ToM experiment reported. YA = Young adults. OA = Older adults. SMD = Standardised mean difference. *d* = Cohen's d. CI = Confidence interval. FB = False belief. NR = Not reported. DV = Dependent variable. Positive effect sizes (i.e., > 0) indicate poorer OA ToM performance relative to YA; negative effect sizes indicate the opposite (i.e., poorer YA ToM).

Young vs. Older Adults

Both individual-level study data and overall group-level analyses were modelled with a random effects model. A weighted effect for each individual study was computed and these were then aggregated to create the overall young vs. old analysis. The data were organised such that a positive effect indicates poorer OA ToM performance (or better YA performance), and a negative effect indicates better OA performance.

All 28 studies contributed to the analysis of the overall YA vs. OA effect. Where studies reported multiple ToM outcomes, effects were pooled to create a single effect size. Overall, in comparison to their younger counterparts, OA performed more poorly on ToM measures (Random effects model: $d = .80$ [.54 – 1.06 95% CI], $Z = 6.01$, $p < .001$), corresponding to a large effect size (Cohen, 1962). Individual study effects and the overall meta-analytic outcome can be seen in the forest plot below (Figure 2.3).

Study or Subgroup	YA Mean	YA SD	YA Total	OA Mean	OA SD	OA Total	Weight	Std. Mean Difference IV, Random, 95% CI
Baksh 2020	21.9	2	30	20	1.88	31	3.7%	0.97 [0.43, 1.50]
Bernstein 2011	7.5	4.25	37	3.3	4.25	37	3.8%	0.98 [0.49, 1.46]
Bottiroli 2016	6.4	0.6	20	4.69	1.69	20	3.3%	1.32 [0.63, 2.01]
Calso 2019	0.79	0.12	30	0.67	0.23	20	3.6%	0.69 [0.10, 1.27]
Calso 2020	19.86	1.46	35	15.77	2.67	30	3.5%	1.92 [1.32, 2.51]
Cavallini 2013	90.55	11.65	30	53.07	15.82	29	3.3%	2.67 [1.96, 3.38]
Duval 2011	96.79	5.75	25	95.23	8.43	25	3.6%	0.21 [-0.34, 0.77]
El Haj 2016	94.99	8.51	40	84.63	15.42	36	3.8%	0.84 [0.37, 1.31]
Fischer 2017	0.89	0.097	86	0.77	0.15	85	4.1%	0.95 [0.63, 1.26]
Franco 2013	8.34	1.95	83	8.23	2.29	89	4.1%	0.05 [-0.25, 0.35]
Grainger 2018	72.06	16.8	51	71.5	16.18	50	4.0%	0.03 [-0.36, 0.42]
Grainger 2019	89.87	5.05	48	79.58	6.48	50	3.8%	1.75 [1.28, 2.22]
Grainger 2020	70.01	15.93	40	64.65	13.61	40	3.9%	0.36 [-0.08, 0.80]
Happe 1998	12.8	2	67	14.9	1.2	19	3.7%	-1.12 [-1.66, -0.58]
Hughes 2019	0.07	0.12	40	0.01	0.13	35	3.8%	0.48 [0.02, 0.94]
Lecce 2019	79.48	19.63	30	57.93	15.95	31	3.6%	1.19 [0.64, 1.74]
Li 2013	10.86	1.58	28	11.43	0.98	28	3.7%	-0.43 [-0.96, 0.10]
Maylor 2002	7.71	1.24	30	4.03	2.24	30	3.5%	2.01 [1.38, 2.63]
Moran 2012	0.94	0.12	29	0.85	0.16	13	3.4%	0.66 [-0.01, 1.33]
Nolaker 2018	16.6	2.65	20	13.88	3.18	19	3.4%	0.91 [0.25, 1.58]
Phillips 2011	87.5	9.31	52	76.4	12.24	36	3.8%	1.04 [0.58, 1.49]
Rakoczy 2012	13.85	1.46	27	12.4	2.16	20	3.5%	0.80 [0.19, 1.40]
Rakoczy 2018	0.95	0.08	40	0.88	0.14	40	3.8%	0.61 [0.16, 1.06]
Reiter 2017	0.74	0.12	55	0.6	0.11	52	3.9%	1.21 [0.79, 1.62]
Slessor 2007	78.13	13.03	40	75.21	14.56	40	3.9%	0.21 [-0.23, 0.65]
Sullivan 2004	9.5	1.3	24	7	2.2	24	3.5%	1.36 [0.73, 1.99]
Zhang 2018	0.67	0.1	87	0.6	0.16	84	4.1%	0.52 [0.22, 0.83]
Total (95% CI)			**1124**			**1013**	**100.0%**	**0.80 [0.54, 1.06]**

Heterogeneity: $Tau^2 = 0.41$; $Chi^2 = 208.20$, df = 26 ($P < 0.00001$); $I^2 = 88\%$
Test for overall effect: Z = 6.01 ($P < 0.00001$)

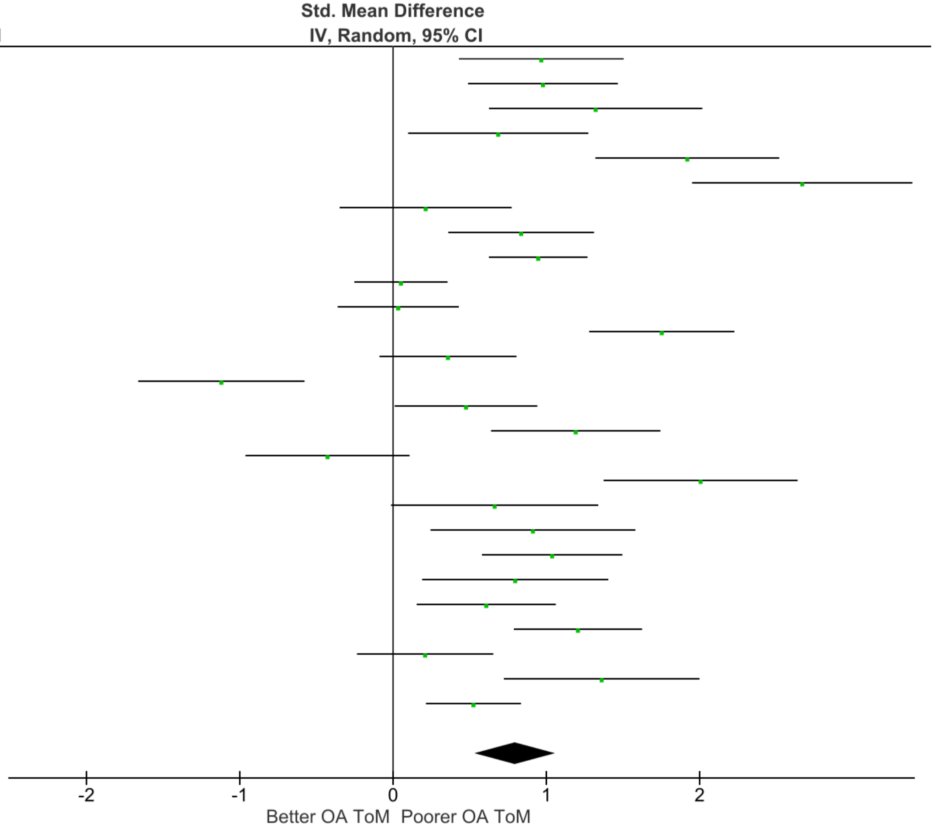


Figure 2.3. Forest plot of studies included in meta-analysis. Studies listed alphabetically by first author surname. Effect sizes for each study are indicated by the position of each green square on the horizontal axis. The horizontal bars indicate the 95% confidence interval for each effect size. The black diamond represents the point estimate and confidence intervals averaged across all individual studies. OA = older adults. YA = younger adults.

Younger vs. Older Adults: ToM Tasks and Participant Characteristics

To understand the specificity of the age effect a series of sub-group analyses were performed on different types of ToM tasks and participant characteristics. Individual, un–aggregated experiment-level data contributed to the sub-group analyses. Task distinctions and categories were adapted from Henry et al. (2013). Tests of sub-group differences were also performed to illustrate between-category differences.

First, the distinction between verbal and non-verbal ToM tasks was examined. Verbal tasks primarily constitute a written or spoken component (e.g., Happé's Strange Stories) whereas non-verbal tasks typically present ToM stimuli visually, such as movies, comic strips, or cartoons (e.g., Wimmer & Perner's (1983) object-location/displacement task). It should be noted, however, many tasks do not adhere to this demarcation of exclusively either verbal or non-verbal as many contain both written and visual elements. For example, the visual Sally-Ann-style false belief task in Duval et al. (2011) comprised a classic change-of-location paradigm that was presented with short written descriptions beneath the stimuli. In the present study, a verbal or non-verbal distinction was made based on the critical ToM task component, and in the example of the above-mentioned Duval et al. (2011) paper, this was determined to be non-verbal.

There were a total of 18 datasets that were categorised as verbal and 17 as non-verbal (a full list of categories and their constituent datasets can be found in Appendix A). For verbal tasks, there was a medium effect (d = .61) in the difference of ToM competence between YA and OA. For non-verbal tasks, there was a substantially (and significantly) greater age effect (d = 1.05). Table 2.3 below includes the results of all sub-group analyses.

Table 2.3.
Statistical outcomes of sub-group analyses.

Taxonomy	Variant	YA vs. OA effect				Test of subgroup differences: X^2, *p*
		No. of experiments	*d* (95% CI)	*z*	*p*	
Domain						
	Verbal	18	.61 (.31 - .91)	3.98	<.0001	$X^2 = 4.11$, $p = .04$
	Non-verbal	17	1.05 (.75 – 1.34)	7.00	<.0001	
Task Taxonomy‡ (V = Verbal, NV = Non-verbal)						
	Movies (NV)	9	1.34 (.92 – 1.77)	6.18	<.0001	$X^2 = 11.60$, $p = .02$
	Static Cartoons-Comics (NV)	6	.96 (.61 – 1.31)	5.42	<.0001	
	Immersive – VR (NV)	1	.47 (.02 – .91)	2.05	.040	
	Vignettes – False belief (V)	16	.63 (.28 – .98)	3.49	<.0001	
	Vignettes – Faux pas (V)	3	.48 (-.13 – 1.09)	1.55	.120	
OA Age†						
	60 – 74	24	.63 (.40 – .86)	4.35	<.0001	$X^2 = 6.12$, $p = .01$
	≥ 75	11	1.27 (.82 – 1.71)	6.52	<.0001	
OA Education^						
	<14 years	14	1.13 (.70 – 1.56)	5.16	<.0001	$X^2 = 3.19$, $p = .07$
	≥ 14 years	10	.58 (.15 – 1.00)	2.64	.008	

(*Table continues*)

Taxonomy	Variant	YA vs. OA effect *No. of experiments*	 *d* (95% CI)	 *z*	 *p*	Test of subgroup differences: X^2, *p*
Domain x OA Education^						
	Verbal, <14 years	7	1.07 (.45 – 1.69)	3.38	<.001	X^2 = 16.47, *p* = .02
	Verbal, ≥ 14 years	7	.20 (-.40 – .81)	.66	.51	
	Non-verbal, <14 years	6	1.19 (.57 – 1.81)	3.76	<.001	
	Non-verbal, ≥ 14 years	4	1.11 (.65 – 1.58)	4.69	<.0001	

Note: ‡ Task taxonomy adapted from Henry et al. (2013). † Based on the average OA' reported mean age. ^ Education only reported in 24 (/35) datasets. YA = Young adults. OA = Older adults. d = Cohen's d. CI = Confidence interval.

A test of sub-group differences – where the null hypothesis is that the effects share the same distribution – showed that these two effects (Verbal vs. Non-verbal) were significantly different (p = .04). In other words, the deficit in OA ToM performance (relative to YA) is greater in non-verbal ToM measures compared to verbal measures.

The specificity of the age effect was also examined across the different types of ToM tasks, namely: Movies (d = 1.34), Static Cartoons-Comics (d = .96), Immersive-Virtual (d = .47, Vignettes False Belief (d = .63), and Vignettes Faux Pas (d = .48). The effect of age was statistically significant across all categories save for Vignettes Faux Pas, which did not yield a significant effect (p = .120). It should be noted, however, that there was only one dataset that constituted the analysis for the Immersive-Virtual task category, and thus the age effect here – though significant – should be interpreted with caution.

Next, the effect was examined as a function of the mean reported age of the OA cohort in the included studies. There was wide range in the mean OA age across the included studies, with the lowest reporting an average OA age of 65 (Zhang et al., 2018) and the highest reporting an average of 84 (Calso et al., 2020). Given the general negative association between social cognitive abilities and age (e.g., see Laillier et al., 2019), it was important to investigate the overall effect within distinct OA age categories. To achieve this, a median-split (median OA age: 73.65) was performed on the reported mean age of each dataset; this led to the creation of two discrete age categories: 60 – 74 and ≥ 75. The YA vs. OA age effect was then probed within each of these two categories. There was a medium age effect within the 60 – 74 grouping (d = .63) and a large effect in the ≥ 75 (d = 1.27). Interestingly, there was a statistically significant sub-group difference (p = .01) which suggests that the difference in YA-OA ToM performance is more pronounced when the OA cohort were, on average, aged 75 and over.

A similar procedure was undertaken with the OA education data (indexed in years), where again a median-split was performed to create dichotomous categories: <14 years and ≥ 14 years. Education was considered an important variable to investigate as it has been linked with individual differences in performance on a wide range of neuropsychological tests in cognitive ageing, some of which are relevant for a functioning ToM. A number of studies were excluded from this particular analysis due to either not reporting education data or reporting education as a categorical variable (see Appendix B for list of excluded studies from education sub-group analysis). The effect of age was significant in both the <14 ($d = 1.13$) and ≥ 14 ($d = .58$) categories. While there was a seemingly larger age effect in the <14 years education category, the test of sub-group differences was not statistically significant ($p = .07$).

Finally, ToM task modality and OA education variables were combined to interrogate the specificity of the age effect on a task modality-by-education level. The justification for this particular demarcation of the data is that we suspected that the specificity of the overall ToM ageing effect may have been confounded in the above-mentioned education sub-analysis. It is reasonable to assume that verbal tasks – which are primarily administered in textual form – are reliant on wider cognitive abilities related to reading and language comprehension; abilities that are not typically purported to be necessary for an operational ToM. Relatedly, language abilities (production and comprehension) in non-pathological ageing have also been shown to be associated with education (e.g., see Thow et al., 2018). With the view to disentangle these features in the context of comparing young vs. old cognitive ToM, we ran sub-group analyses on ToM domain (verbal, non-verbal) as a function of OA education (<14 years, ≥ 14 years);

which resulted in the creation of four cells: 1) Verbal, <14 years; 2) Verbal, ≥14 years; 3) Non-verbal, <14 years; and 4) Non-verbal, ≥14 years.

Three of the four cells yielded effects that were statistically significant, except for Verbal, ≥ 14 years, where there was no significant age effect (d = .20, p = .51). For brevity, these effects are not enumerated here, but can be seen in Table 2.3. The test of sub-group differences showed that the magnitude of the effects from each cell were significantly different, suggesting that the variability in the magnitude of the age effect may be contingent on the domain of the ToM task *combined* with the education level of the OA cohort.

Heterogeneity

Heterogeneity was assessed via the conventional method using Cochran's Q statistic and I^2. Cochran's Q – distributed as a chi-square statistic – is the classic statistical test of effect size inconsistency and is computed as the weighted sum of squared differences between study effects and the pooled effect across studies. As a complement to the Q statistic, the I^2 value represents the percentage of variation between studies that is a result of innate heterogeneity rather than chance, with 100% representing complete and full variation between studies due to *true* effect size inconsistency and 0% suggesting that all variation between studies is due to chance alone (Higgins et al., 2003). Together, the Q and I^2 illustrate the statistical significance and magnitude of heterogeneity across the meta-analysed effects.

In the overall age effect analysis, there was substantial and significant heterogeneity between the included studies ($Q[X^2] = 208.20$, $p < .001$, $I^2 = 88\%$), suggesting that 88 per cent of the variation between the reviewed effect sizes was down to innate inconsistency (as opposed to chance); this equates to “considerable” heterogeneity according the conventional rule-of-thumb (Schünemann et al., 2013). To probe the issue of heterogeneity further, first, statistically anomalous results were examined.

The effect of one study (Happé et al., 1998) was three SDs below the overall pooled effect (d = -1.12 [95% CI = -1.66 – -.58), and was in fact the only included study to report a statistically significant *negative* effect, where OA exhibited better ToM performance compared to their younger counterparts. To examine the relative contribution of this outlier on the measure of heterogeneity and indeed the overall age effect, the core YA vs. OA meta-analysis was re-run without Happé et al., (1998). This marginally decreased heterogeneity by four percentage points ($I^2 = 84\%$). Relatedly, removal of Happé et al. also subsequently increased the overall meta-analytic effect of age (New model: d = 0.86 [0.66 – 1.06 95% CI], Z = 8.41, $p < .001$; Original model: d = .80 [.54 – 1.06 95% CI], Z = 6.01, $p < .001$).

Sub-group analysis is a valuable tool in examining the overall meta-analytic effect within a smaller, theoretically-driven selection of conceptually homogenous experiments. As the overall effect is probed within and across specialised ‘groups’, more nuanced and refined conclusions can be drawn. In the context of effect inconsistency, as studies are classified and clustered according to conceptual likeness, an analysis of sub-groups allows researchers to examine the relative change in heterogeneity, with the view of reducing effect inconsistency by carefully constructing qualitatively ‘valid’ and conceptually-driven sub-groups. Practically, a new and separate set of heterogeneity statistics are computed for every sub-group classification. In the current analysis, however, though heterogeneity reduced with the addition of sub-groups, the minimum I^2 value was 74% which still amounts to “substantial” heterogeneity.

It should be noted, however, large variation or inconsistency in effects should not be interpreted as a commentary on the quality of the evidence reviewed (Schünemann et al., 2013). It is instead perhaps a better reflection of inhomogeneity in task design, within and cross-study participant variability, differences in the calculation of outcome measures, and other methodological particulars. Therefore, in the present case, effect inconsistency can be thought of as an indicator of non-standardisation of methodological/experimental design rather than qualitative differences in study *quality*.

Publication Bias

Publication bias was examined through a funnel plot (Figure 2.4). To determine the degree of bias, the extent of plot asymmetry was examined. Where there is no bias, the scatter plot should resemble a symmetrical inverted funnel shape with a wide base (consisting of small studies with large effect estimate variability) and a narrow top (larger studies with small effect estimate variability).

In the present case, though there is slight asymmetry in the shape of the plot, the violation is not sufficiently severe to determine that significant publication bias exists within the included studies of this review. As such, it is reasonable to propose that publication bias was not a significant concern in the current analysis.

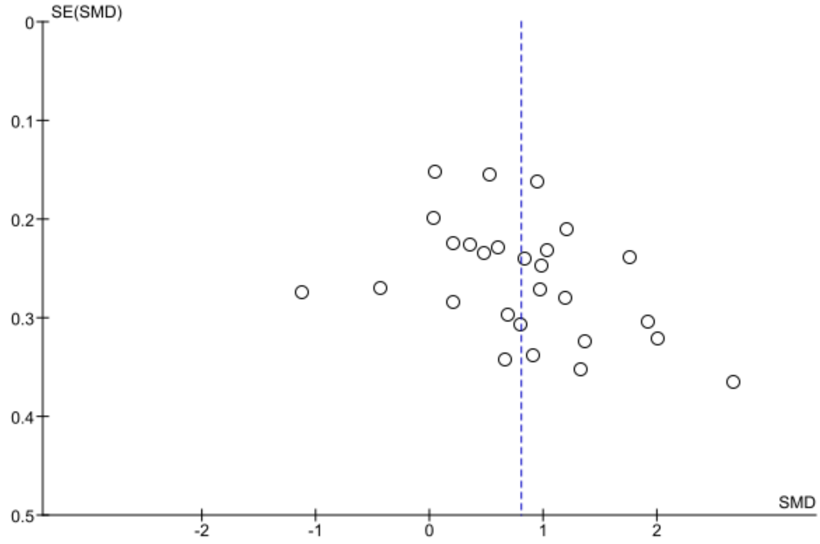


Figure 2.4. Funnel plot. Effect estimates are presented on the *x*-axis and study precision (study size; S.E. of Standardized Mean Difference) is shown on the *y*-axis.

Discussion

This quantitative review aimed to clarify the effect of age on ToM proficiency. A further motivation of the current review – which at the time of writing is the first of its kind – was to examine only measures of cognitive ToM in the context of an ageing effect. Finally, and perhaps most importantly, we aimed to highlight the specificity of an effect of age across the varied and heterogeneous nature of ToM paradigms.

Overall ToM age effect

With the exception of a small number of null and one negative effect (Happé et al., 1998), a strong age effect (d = .80, range -1.12 – 2.67) emerged across all measures of ToM, irrespective of task parameters. This is consistent with previous meta-analyses of ageing ToM (e.g., Henry et al., 2013). In line with studies that have mapped ToM abilities across the adult lifespan (e.g., Laillier et al., 2019), a young vs. older adult comparison of the published literature suggests that OA exhibit poorer ToM compared to YA. While there was significant heterogeneity in effect sizes, the overall effect of age was robust even after taking into consideration different task types and inconsistent sampling.

Specificity of the age effect across task and sampling considerations

Verbal vs. Non-verbal ToM Tasks

We examined the nature of an effect of age across two broad distinctions of ToM tasks, verbal and non-verbal. While there was indeed an overall effect of age within both verbal and non-verbal tasks, a difference in young vs. old ToM was substantially more pronounced in non-verbal paradigms. One reason for this disparity may be the underlying cognitive abilities that subserve each task domain. Verbal ToM tasks are primarily presented in textual form, and therefore ostensibly tap language abilities (e.g., reading and comprehension) more so than non-verbal tasks. With this in mind, in ageing, while lexical retrieval is generally impaired, vocabulary or lexical diversity has been found to be preserved and even improved in older age (Kemper & Sumner, 2001; McCabe et al., 2010). It is plausible, therefore, while not completely abolished, the effect of age in verbal tasks is in part attenuated (relative to non-verbal tasks) due to older adults' enhanced capabilities in aspects of language. Indeed, in preschool samples, language training interventions have been shown to improve performance across a range of ToM tasks in pre-post experimental designs (e.g., see Hale & Tager-Flusberg, 2003). However, to empirically measure the association between OA' language abilities and performance in verbal vs non-verbal ToM tasks, studies must examine the specific relationship between language skills and ToM performance in advanced age. This would disentangle the influence of language in verbal ToM tasks and provide a more refined estimate of an age-related effect.

Task Type

To further investigate the observed heterogeneity in the overall YA vs OA comparison, tasks were grouped into one of five classifications: *movies, static cartoons-comics, immersive – VR, vignettes false belief,* and *vignettes faux pas*. These classifications were adapted from Henry et al. (2013) and Quesque and Rossetti (2020). We found that there was substantial variation in the magnitude of the effect size across these classifications, with the greatest effect seen in the movies classification and the smallest effect seen in the vignettes faux pas grouping, where the effect of age was statistically non-significant.

This variability in effects perhaps speaks to the differential demands that are involved behind each task type, where again tasks that were non-verbal in nature – i.e., movies and static cartoons-comics – produced greater age effects relative to the two verbal tasks: vignettes false belief and vignettes faux pas. The exception here, however, is the immersive VR classification which produced an effect smaller in magnitude than vignettes false belief; though, the immersive VR grouping only contained a single study and therefore it is difficult to draw meaningful comparisons with the other classifications.

One interpretation as to why movies produced the largest effect is perhaps due to increased processing and executive demands from task manipulations that are naturally less carefully controlled relative to, say, static picture-based paradigms. For example, the MASC paradigm is reported to be a naturalistic and sensitive measure of ToM, comprising complex social scenarios depicting interpersonal relationships between four different characters, who each have their own comprehensive backstories. The scenes are rich in detail and the characters were composed with basic needs and sub-needs that varied throughout the task. The authors purposely built in ambiguity between characters' basic and sub-needs and created "[Items] that pose particularly challenging demands on social cognitive functioning" (Dziobek et al., 2006, pp. 627). To illustrate the potential difference in processing demands between tasks, contrast the MASC paradigm with the Geometric Triangles Animation task utilised in Zhang et al. (2018). Based on Castelli et al.'s (2002) original animated triangle task, participants are shown two animated, non-human inanimate shape forms that move in a manner suggesting intentionality (pretending, deceiving, etc.), and participants are simply tasked with discerning the intentional state of each shape. Both of these paradigms can be broadly categorised as movie- or film-based, as they are dynamic 'motion picture' tasks.

However, there are clear and obvious qualitative differences in the difficulty or cognitive demands of each task, where the MASC perhaps has a greater reliance on faculties of working memory and attention (to manage multiple characters, backstories, and needs) in comparison to the (potentially) less cognitively burdensome animated shapes tasks. This variability in task complexity, and therefore task demands, may in part explain the substantial heterogeneity between *and* within task types with respect to an effect of age in ToM.

The Effect of Age, Sampling and Task Considerations

In addition to effect variability being driven by inconsistencies in task complexity, there were also noticeable differences in the demographics (age and education) of OA samples across the included studies of this review. With respect to the average OA reported age, studies varied from around 65 in Zhang et al. (2018) to around 85 in Calso et al. (2020). This variability in age is an important consideration when examining the nature of an age-related difference in ToM abilities as studies have shown a general downward trajectory of SCF across the adult lifespan (Laillier et al., 2019). In line with this view, we found a substantially greater effect of age from our dichotomised age group analysis, where studies whose OA sample, were, on average, aged 75 and over, exhibited stronger age-related effects relative to studies whose OAs were aged 74 or less.

There was also substantial variability in the years of education across OA samples; where some studies reported as little as 9.8 years of education (Calso et al., 2019), others reported almost 16 years (Li et al., 2013). In a specific sub-group analysis, where we dichotomised level of education through a median-split, there was some differentiation in the effect of age between the two education categories: <14 years education and >14 years. It should be noted, however, this sub-group analysis did not reach statistical significance (p = .07), and any subsequent conclusions based on education from this analysis should be interpreted with this context in mind. Education did, however, interact with task domain and is discussed later in the Chapter.

The variability in levels of education is an important consideration as education has been related to composite measures of intelligence (fluid and crystallised) in later life (see meta-analysis by Ritchie & Tucker-Drob, 2018). Relatedly, intelligence (in particular, fluid intelligence) has been found to partially explain individual differences in ToM abilities (Ibanez et al., 2013). While there has been no dedicated, objective investigation of the link between education and ToM abilities in later life, given the protective quality of higher levels of education on wide ranging measures of cognitive functioning, some of which are fundamental to an operational ToM, it is plausible to assume that the variability in OA education, coupled with particular task variations across studies, may in part explain the inconsistency of the overall age effect.

We suspected that there might be an interplay between task domain and education as it is reasonable to assume that task difficulty/characteristics may interact with education/intelligence. It is worth noting, however, we did not perform an objective statistical *interactive* test to examine the potential modifying effect of one variable on the other. We simply combined the Task Domain and OA Education variables to create four distinct cells: *Verbal <14 years, Verbal ≥ 14 years, Non-verbal <14 years*, and *Non-verbal ≥ 14 years*; and examined the difference in effect sizes between these new, composite classifications. As expected, we found that the effect of age varied according to domain and OA' level of education, where the greatest age effect was seen in the *Non-verbal, <14 years* group (d = 1.19) and the smallest effect, which was also non-significant, in the *Verbal ≥ 14 years* group (d = .20). In our previous, individual sub-group analyses of Verbal vs. Non-verbal tasks and <14 years vs. ≥ 14 years of education, we showed that the effect of age was smallest in non-verbal tasks and when the OA samples were highly educated. Logically, the combination of these two classifications has led to the effect of age being abolished altogether. This again confirms our suspicion of the effect of age being contingent upon methodological intricacies, and that the classic ageing ToM literature may have perhaps inflated an effect of age by introducing task- and sample-related artefacts.

Difficulties in Creating Task Taxonomies

It is worth noting that while the taxonomies used in this study were based on previous reviews, some classifications were arbitrary and somewhat contrived. Many tasks straddled multiple classifications, including domain and even task type. Consider the MASC, for example. While the main component was presented in a film format, follow-up questions that probed participant understanding of mental states, and on which the outcome measures were based, were presented textually. While we in the current review classified investigations that used the MASC as *Non-verbal*, on the basis that the critical social 'scene-setting' is presented through the medium of film, it is equally arguable that the MASC could be categorised as 'Verbal', or at least 'Mixed' (a 'Mixed' classification has been used in previous ageing ToM meta-analyses, for example, see Henry et al., 2013). Our quantitative analysis and subsequent discussion should be interpreted with such a context in mind.

Heterogeneity

In the overall effect of age, which meta-analysed the outcomes of 28 studies (with 35 datasets, $n = 2137$), there was substantial and large heterogeneity both objectively and subjectively in terms of the nature of tasks deployed. While the effects became more homogenous in the sub-group analyses when studies were clustered according to domain, modality, and participant characteristics, there was still significant heterogeneity in the observed effects. No sub-group analyses yielded non-significant heterogeneity in effect size estimates.

As we have discussed above, the effect of age seems to fluctuate depending on a range of methodological features. There was variability in the nature of tasks deployed, where some were presented in movie format, and others in static picture-based formats, some presented a single protagonist whereas others presented multiple, among other variabilities. Even within the same task taxonomy, there was variability in the cognitive or processing demand underlying ToM/social processing. While statistical heterogeneity in meta-analyses is not an indicator of study *quality*, the core takeaway here is that the reported effect of age varies substantially across the published ToM literature, and this effect inconsistency remains even after studies are sub-divided according to methodological and sampling likenesses.

Limitations & Future Directions

There are a few limitations in this study. First, the quantitative analysis only included data derived from the ToM (or mentalizing) component from each individual study. The current analysis did not take into consideration data from non-ToM control conditions. Accordingly, it is not possible to explore and identify whether the overall age effect remains the same or if it is modified in a matched control condition. Henry et al. (2013) included both ToM and non-ToM matched control tasks in their meta-analysis and found that the magnitude of the age effect decreased in the non-ToM condition, which in their view, is again evidence of an age difference in ToM.

Second, there was no objective account of executive function or fluid intelligence in the current analysis. Indeed, the majority of the included studies reported complementary measures of cognitive ability (e.g., working memory, inhibitory control, task switching, processing speed, etc.) and others adjusted the between-groups analyses to account for differences in these capabilities. There is a well-established effect of age on a range of cognitive domains and neurocognitive indices (see Salthouse, 2010, for a review on cognitive ageing); and an operational ToM, especially in ageing, is thought to rely on intact executive functions (see Chapter 3). Therefore, future reviews should incorporate measures of cognitive function (beyond social cognition) to examine the interplay between EF, ToM, and ageing.

In addition to incorporating measures of EF on a broader level, a more fine-grained task analysis of ToM paradigms and their associated executive demands would be advantageous in furthering our understanding of the nature of age-related alteration of ToM. Conventional paradigms often confound ‘social’ and ‘non-social’ manipulations such that it is difficult to disentangle and delineate age-related social cognitive decline from the more domain-general, slowing that is common in advanced age. We know that executive abilities decline with age (Salthouse, 2019), and therefore conventional ToM protocols may inadvertently inflate age-related effects through non-social, incidental task demands. For a more nuanced view of an age effect and how it interacts with domain-general demands of ToM tasks, future meta-analyses should group tasks according to non-social cognitive demands to examine whether the reported effect of age is maintained or modified according to cognitive load. Though, this is difficult given between- and within-task variability.

Another consideration is inconsistent definition of 'older adult' or 'ageing' in the literature. As we have seen, there is great variability in the reported OA average age across studies, and where studies have created young-old and old-old age demarcations, there is again variability in the cut-offs used to separate these groupings. With the aim to create a more standardised and comparable ToM ageing literature, we propose that ageing is perhaps best thought of as starting at 60 years, and relatedly, young-old comprising 60 – 75 years, and old-old comprising 75+ years. It is hoped that with a more standardised grouping of age, general cognitive slowing in advanced age can be more carefully controlled and managed and a more refined account of age-related effects on social cognitive function can be attained.

Finally, the *quality* of the evidence was not evaluated in this review. This is an important consideration as quality assessment ensures that the individual studies included in the review are evaluated according to standardised criteria with respect to methodology and statistical approach, such as reducing/eliminating sampling bias, participant randomisation, condition randomisation, correcting for multiple comparisons and family-wise error, and so on. Without evaluating the quality of the evidence that feeds into the overall meta-analytic effect, the conclusions of the meta-analysis may not be as robust and hard-wearing. Though, it should be noted that previous quantitative reviews of ToM in ageing (e.g., Henry et al., 2013) have not included quality assessments.

Conclusion

Social-cognitive abilities, and in particular ToM, are important for everyday social interactions and interpersonal relationships. While the early literature on ageing ToM produced somewhat contradictory findings, contemporary work has suggested that there is a general deterioration of ToM in advanced age. In our review of the evidence, we found that ToM competence does indeed seem to diminish with age, but importantly, the magnitude of this decline is modified – and in some cases, abolished – depending on task and sample considerations. It is our view, then, due to the large statistical heterogeneity between the studies included in the analysis and in methodological designs of the tasks employed, with different types of tasks tapping separate and distinct executive demands, there is likely an overestimation of the age effect. Still, the results of this review have important implications for future investigations of ageing ToM, and we recommend that researchers carefully construct tasks in a way that minimizes (or at least controls for) cognitive load outside of what is ordinarily required to mentalize and reason about the beliefs of others.

Chapter Three: Sources of Conflict in Ageing ToM

Sources of cognitive conflict and their relevance to Theory of Mind proficiency in healthy ageing.

An amended version of this chapter has been accepted for publication at *Psychological Science*. An earlier version of this chapter was also uploaded as a preprint onto PsyArxiv: https://psyarxiv.com/mzv5u/. This study was formally preregistered following an Open Science Framework protocol, which is available from DOI: 10.17605/OSF.IO/DC8CE.

Chapter précis and background

In the meta-analysis in Chapter 2, we showed that the effect of age on ToM competence is highly heterogeneous, with the magnitude of an age effect varying according to both task and sampling considerations. In some instances, the effect of age was abolished. One particular reason for this heterogeneity is perhaps down to the qualitative differences in task difficulty, where some ToM paradigms have greater non-social cognitive loads than others. In typical false belief tasks, successful performance relies on both social and non-social (or executive, domain-general) processing. Non-social demands rely on intact executive functions, which are also known to decline with age. However, within the ageing false belief literature, this is yet to be disentangled; processing that is non-social in nature is often confounded with social cognitive function. To address this, we assessed how older (n=50, aged 60-79) versus younger healthy adults (n=50, aged 18-29) were affected by three theoretically relevant sources of conflict within ToM: competing Self-Other perspectives; competing cued locations, and outcome knowledge. We examined which best accounted for age-related difficulty with ToM. Our data show unexpected similarity between age groups when representing a belief incongruent with one's own. Individual differences in attention and response speed best explained the degree of conflict experienced through incompatible Self-Other perspectives. However, OAs were disproportionately affected by managing conflict between cued locations. Age and spatial working memory were most relevant for predicting the magnitude of conflict elicited by conflicting cued locations. We suggest that previous studies may have underestimated OA's ToM proficiency by including unnecessary conflict in ToM tasks.

Introduction

Previous studies of normal ageing and 'Theory of Mind' (ToM), the ability to infer another person's thoughts, beliefs and desires, have produced contradictory findings regarding whether ToM competence declines in older adulthood (Henry et al., 2013; Love, 2015; Phillips et al., 2011). Methodological limitations, however, through some tasks making excess demands on executive function (EF), may explain this disparity (Love, 2015). EF and processing speed typically deteriorate with healthy ageing (Salthouse, 1996; 2010; 2012), but cognitive conflict may also be embedded within representing other peoples' cognitive perspectives (Austin, Groppe, & Elsner, 2014; Leslie, Friedman, & German, 2004), making understanding age-related differences in ToM proficiency difficult.

EF is important for an operational ToM (Austin, Groppe, & Elsner, 2014; Vetter et al., 2013). ToM often involves reasoning about others' beliefs that may differ from our own. Sometimes there is more than one other person and our own knowledge of the right answer may vary in certainty. These factors are often confounded in the classic ToM literature. Working memory, attention, and inhibition have been suggested to support mental state representation through managing conflict (Austin, Groppe, & Elsner, 2014; Leslie, Friedman, & German, 2004). Such conflict could arise from competing cued information, where attentional resources must be disengaged from one information source to select another – a typical feature of false belief (FB) paradigms which are used to assess ToM. Likewise, knowledge of an event's outcome may also interfere with an individual's ability to reason, due to bias towards one's own, salient self-knowledge – termed a 'curse of knowledge' (CofK; Birch & Bloom, 2004; 2007; for a review, see Ghrear, Birch, & Bernstein, 2016). To make predictions based on an agent's FB, one must inhibit one's own perspective to adopt the other person's. Indeed, in healthy adults, false- versus true-belief reasoning is associated with slower, more error-prone behavioural performance (Apperly et al., 2008; Apperly et al., 2011). Manipulation of core parameters within the FB task demonstrates additional processing associated with ‘self-perspective inhibition’, suggesting that incongruent Self-Other cognitive perspectives may create conflict that is distinct from other sources of conflict within ToM tasks (Hartwright et al., 2015; Samson et al., 2005; Samson et al., 2015). It is not, however, clear whether having a privileged knowledge of reality, or the mismatch in the cognitive perspectives of Self and Other, is the basis of competition in FB reasoning. Moreover, it is unclear whether the source of competition may explain conflicting findings in healthy ageing.

In this study, we examined which factors are responsible for variation in the difficulty of reasoning about the beliefs of others across age groups. We assessed three theoretically relevant potential sources of conflict in ToM: privileged outcome knowledge, congruence of Self-Other perspectives and competing cued locations. Research assessing the CofK suggests that one’s own, privileged knowledge can cause interference when judging what others know (Birch & Bloom, 2004; 2007). Further, when that knowledge differs between Self and Other, the conflicting Self perspective must be inhibited (Hartwright et al., 2015; Samson et al., 2005; Samson et al., 2015). On this basis, we tested how a participant’s knowledge of reality, and incongruence between the participant’s “Self-perspective” and an agent’s “Other-perspective” contribute to processing difficulty. These two aspects are confounded in the classic object transfer FB task, consequently, the present study aims to disentangle these as candidate sources of conflict. Furthermore, giving the correct answer in FB paradigms often requires participants to shift attention between competing locations, typically cued by a participant’s representations of where the object is located and where the other person thinks it is located (see Friedman & Leslie, 2005). This is a feature of many FB paradigms though, unlike differences between Self- and Other-perspectives, is not an essential feature of FB problems. In this study, we de-confounded this factor from effects of knowledge-of-reality in conditions where alternative locations corresponded to the beliefs of two different agents. Building on research suggesting that incongruent Self-Other Perspectives create conflict distinct from other sources of conflict within ToM (Hartwright et al., 2015; Samson et al., 2005; Samson et al., 2015), we hypothesised that there would be greater cognitive effort associated with holding in mind competing Self versus Other perspectives versus managing alternate cued locations.

Furthermore, we aimed to understand how these three sources of conflict are relevant to ageing in ToM. When compared with younger adults, older adults demonstrate greater hindsight bias when informed with outcome knowledge (Bernstein et al., 2011), more difficulty with managing incongruence between beliefs, and larger biases towards cued locations in FB reasoning (Bernstein, Thornton, & Sommerville, 2011). Older adults might therefore experience difficulty with self-perspective inhibition, attending and managing conflict from multiple cued locations, and handling incongruence between beliefs – all aspects pertinent to FB representation, but not all essential to ToM. Mental-state representation has consistently been shown to recruit different brain systems to non-mental representation (Saxe & Kanwisher, 2003; Saxe & Powell, 2006), but those neural systems for ToM interact with systems for EF (Hartwright et al., 2012; 2013; 2016; Mars et al., 2012). Dwindling underlying baseline connectivity of brain regions typically associated with ToM – particularly the temporoparietal junction (TPJ) – only partially predicts poorer performance in older adults (Hughes et al., 2019). Given that prior research shows differentiation between younger and older adults in ToM as a function of EF demands (Bailey & Henry, 2008;

Bottiroli et al., 2016; German & Hehman, 2006), it is important to better understand how conflict affects ToM processing more broadly and in ageing. We therefore assess, is there a psychologically relevant age decline in managing competing Self-Other perspectives, or are older adults disproportionately affected by methodological confounds, like the CofK and the need to manage competing cued locations? By manipulating psychologically relevant parameters within a single ToM task, we evaluate which cognitive components associated with belief-reasoning explain age-specific deficits in performance. We also used standardised measures of EF to predict the magnitude of conflict elicited by our ToM manipulations, to further explore the neuropsychological bases of our results.

Materials and Methods

Participants

One hundred and two adults with no self-reported neuro-psychiatric history and normal or corrected-to-normal vision participated in the study. Two participants' data were excluded: one younger adult (YA) due to a methodological issue and one older adult (OA) for scoring beyond the cut-off point in a dementia screening measure. Thus, the final sample comprised 100 participants; 50 YA (21 male; age range 18-29 years, mean age = 20.2) and 50 OA (18 male; age range 60-79 years, mean age = 67.9). To be consistent with the literature, we defined 'young' participants as individuals aged 18 – 29, and 'older' as age 60 – 79. This is in keeping with previous ageing ToM studies. YA were recruited via the university's research participation scheme, university noticeboards and email advertisements to staff; OA were recruited from the university's research panel, local interest and hobby groups, university noticeboards and email advertisements to staff. YA were either compensated with course credits or a small honorarium and all OA received a small honorarium for their participation. All older adults were community-dwelling and 52% were educated to at least undergraduate degree-level. The study was approved by Aston University's Life & Health Sciences Ethics Committee (Appendix C). All participants gave written informed consent prior to participation.

The Power Analysis for General Anova designs tool (PANGEA; Westfall, 2015) was used to conduct a post-hoc sensitivity analysis. Given our sample size of 50 per group, with 18 replicates per observation in our primary task, we had ~90% power to detect three-way interactions with a small effect size (Cohen's d = 0.2).

Design

Theory of Mind Abilities: False Belief Task

The current ToM paradigm was based on Apperly et al. (2011) and Hartwright et al. (2012) and was written in E-Prime (v2.0; Psychology Software Tools, Pittsburgh, PA). The current task was non-inferential: participants are explicitly made aware of others' belief states.

The current ToM task consisted of a three-factor (2x2x2) design wherein each factor manipulated whether a theoretically-based potential source of conflict was high or low (indicated with a subscript 1 or 0 respectively); see Table 3.1 below. The first factor, termed 'Knowledge of Reality' (KoR), varied the presence of the participant's explicit knowledge about reality, and was based on prior work suggesting one's own self-knowledge can cause interference when representing that of another. The KoR manipulation resulted in a 'reality unknown' (KoR_0) and a 'reality known' (KoR_1) condition. The second factor, termed 'Other-Other Congruence' (OOC), manipulated the congruence of two agents' perspectives, resulting in a minimal conflict (congruent) and maximal conflict (incongruent) condition (OOC_0 and OOC_1 respectively). The third factor, termed 'Self-Other Congruence' (SOC), concerned the congruence of the participant's and the *target* agent's perspectives, where the presence of conflict between those perspectives was manipulated. As with the OOC condition, this resulted in a minimal (congruent) and maximal (incongruent) conflict condition (SOC_0 and SOC_1 respectively). These latter two factors were based on work suggesting that ToM reasoning is supported by executive selection to resolve competition between salient cues.

Table 3.1.

Summary of Experimental Factors and Levels

		factor levels	
Condition	Factor descriptor	low[a] conflict	high[a] conflict
Knowledge of Reality (KoR)	Manipulates whether the participant is given explicit knowledge about the true state of affairs	reality unknown (KoR_0)	reality known (KoR_1)
Other-Other Congruence (OOC)	Manipulates the congruence of two agents'[b] beliefs about what is the true state of affairs	other-other congruent (OOC_0)	other-other incongruent (OOC_1)
Self-Other Congruence (SOC)	Manipulates the congruence of the participant's and the target agent's[b] beliefs about what is the state of affairs	self-other congruent (SOC_0)	self-other incongruent (SOC_1)

Note. [a]The deemed 'level' of conflict was theoretically driven. [b]The agent(s) beliefs can be true or false as, unlike the participant, the agents have no knowledge of reality.

The three-factor design was formulated into a computer-based task where participants were required to respond from a target agent's perspective (ToM trial), or based on what they, themselves explicitly knew (an anti-strategy trial, herein termed a 'filler'). Each experimental trial outlined a game where a magician hid a ball in one of three cups and subsequently shuffled the cups away from view (see Figure 3.1 below). Participants were required to indicate either i) where a target agent believed the ball was hidden (ToM trial) or ii) where the participant themselves thought the ball was (filler trial). Each trial comprised a sequence where each of the two agents indicated where they thought the ball was hidden, plus a clue – which the agents were not privy to – regarding what was inside one of the three cups. For each trial, a response probe was presented which indicated which of the two agents was the target agent and the nature of the response the participant should give (ToM or filler). The fillers were developed to confirm that participants were attending to the clue regarding where the ball really was, by responding with the true location of the ball as this clue permitted differentiation in beliefs between the participant and the agents. Note that the fillers can be solved without ToM reasoning and therefore were used only to identify and exclude participants who were not attending to the task appropriately.

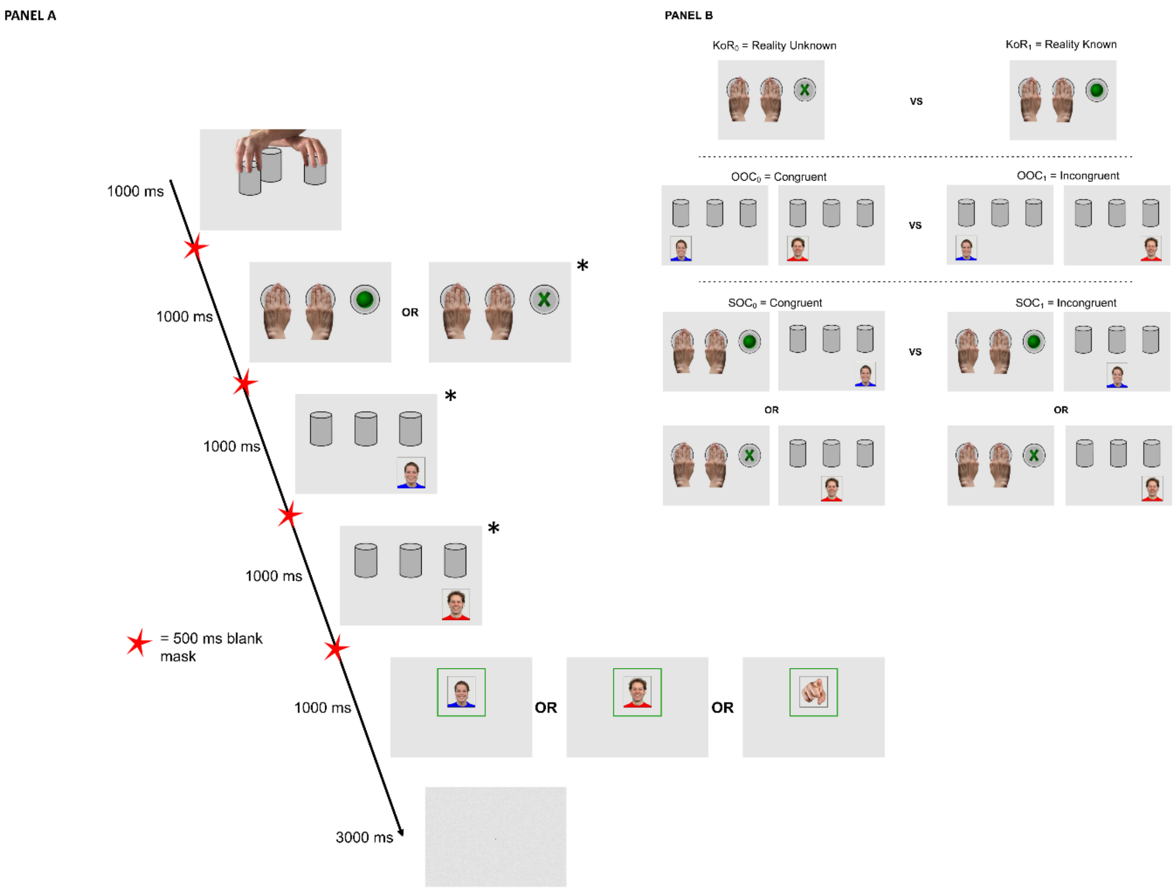


Figure 3.1. Panel A: A single trial outline with event timings. *Order of events varied across trials. Panel B: Schematic of the three experimental factors and how these were achieved. KoR = Knowledge of Reality: reality unknown (KoR_0), where an X was shown to indicate absence of the ball from that location; reality known (KoR_1), where the ball was shown to highlight its true location. OOC = Other-Other Congruence: the two agents' beliefs about the location of the ball were congruent (OOC_0) or incongruent (OOC_1) with one another. SOC = Self-Other Congruence: the target agent's belief about the location of the ball was congruent (SOC_0) or incongruent (SOC_1) with the participant's belief about the location of the ball.

Each trial comprised five static images, followed with a central fixation mark (see Figure 3.1A). The first image always depicted a magician shuffling three cups. Three further images were then presented. The order of presentation of these three images was counterbalanced using a Latin square and randomised. One image showed the magician's hands obscuring the contents of two of the three cups. In the unobscured cup, either a green ball or an X was shown to indicate the presence (green ball) or absence (green X) of the ball, respectively. The participant only ever knew the contents of one cup per trial, consequently, they either knew explicitly the true location of the ball (ball shown), or they had to infer that it was under one of the two obscured cups (X shown). Two further images depicted one of the agents in front of one of the three cups, indicating which cup that agent believed the ball was located in (both agents' beliefs were indicated in every trial). Following presentation of the three images, a response probe was shown. This depicted either an image of one of the two agents, requiring the participant to respond with where that agent thought the ball was (ToM trials), or an image of a hand with a finger pointed toward the participant where – based on the earlier clue – the participants had to respond with where they themselves thought the ball was (filler trials). Participants used number keys 1, 2 and 3 on the number pad of the computer keyboard to indicate their selected cup (left to right, where cup one was coded as 1 on the number pad). The eight experimental conditions were created by manipulating whether the participant knew where the ball was (KoR_1), or if the green X was shown, leaving the location unknown (KoR_0); whether the two agents' beliefs about the location of the ball were congruent (OOC_o) or incongruent (OOC_1); and whether the participant's and target agent's beliefs about the location of the ball were congruent (SOC_0) or incongruent (SOC_1), as outlined in Figure 1B. By varying KoR, OOC and SOC, eight conditions were created: $KoR_0OOC_0SOC_0$; $KoR_1OOC_0SOC_0$; $KoR_0OOC_1SOC_0$;

$KoR_0OOC_0SOC_1$; $KoR_1OOC_1SOC_0$; $KoR_1OOC_0SOC_1$; $KoR_0OOC_1SOC_1$; $KoR_1OOC_1SOC_1$, where each condition described the state of affairs in relation to the target agent, as indicated by the response probe. The study comprised 18 repetitions of each experimental condition for trials and 9 repetitions of each condition for fillers. This resulted in reaction-time and accuracy data for 144 trials of interest and 72 fillers. The number of repetitions of each condition, the location of the ball, and the target agent were counterbalanced across the experiment. Participants completed four counterbalanced experimental blocks, each containing 54 trials. Reaction-times were collected based on the time taken to respond following the onset of the response probe. Accuracy was treated as identifying the correct cup, as required by the response probe. Omissions were treated as errors. Overall, the ToM experiment comprised 216 trials, equally split across 4 blocks (54 per block; block duration = 9 mins; each trial = 10 s). Each block comprised 36 ToM trials and 18 fillers.

Our novel ToM task was non-inferential. While there is some uncertainty in the literature about whether true ToM occurs in the absence of inference, it should be noted that ToM and the inference of others' beliefs, though commonly referred to as interchangeably, are not indistinguishable. Certainly, in everyday social situations others might plainly proclaim what they think or feel, and holding these beliefs in mind while also resisting interference from one's own (potentially) conflicting perspective is cognitively burdensome. Research has shown, for example, after being made aware of someone else's beliefs, both children and adults are still prone to egocentric interference (Apperly, Back, Samson & France, 2008). Here, participants are not required to make inferences about what the other person thinks nor are they required to make inferences about behavioural consequences, yet one's own perspective still seems to interfere when representing the other's (overt) belief. Further, previous neuroimaging work from

Hartwright et al. (2012) has identified increased activation of the ToM network during representation of non-inferential mental states. In the same study, Hartwright and colleagues also employed the well-known ToM Localizer task (Saxe & Kanwisher, 2003) where participants, in one condition, inferred protagonists' mental states to resolve the task. Substantial overlap in the activation of the TPJ was found between the two tasks, Hartwright and colleagues' non-inferential task and Saxe & Kanwisher's inferential tasks. Taken together, these converging lines of evidence suggest that non-inferential and inferential mental state representation tap similar task demands and share equivalent patterns of neural activation. Accordingly, our novel design comprised a non-inferential false belief task.

Furthermore, the current paradigm is justified as a measure of ToM—or more precisely, FB—as it satisfies two fundamental tests of belief inference. According to Quesque and Rossetti (2020), two criteria must be considered when evaluating ToM tasks: mentalizing (inference of belief) and a distinction in self-other mental states. Quesque and Rossetti note that tasks must require participants to hold in mind a belief or intentional state of an external agent, and that this mental state must be separate and distinct from the participant's own beliefs. In the current task, participants were required to mentalize about the (explicit) belief states of one of two external agents, and these belief states were independent of the participant's own—and sometimes, privileged—knowledge of reality. We therefore consider the current paradigm, and its iteration in Chapter 4, to be a faithful measure of ToM reasoning ability.

Theory of Mind Abilities: Self-Reported Perspective Taking Capacity

The Interpersonal Reactivity Index (IRI; Davis, 1980) comprises four self-report subscales: Perspective Taking (PT), Fantasy, Empathic Concern and Personal Distress.

Although we administered the full scale to ensure reliability of the measure, we were primarily interested in data from the PT subscale as this is said to be indicative of a participant's (self-reported) proficiency with taking other people's cognitive perspectives. This measure did not form part of any preregistered hypotheses but was used for exploratory analyses.

Neuropsychological Testing

Participants' EF was evaluated using the Cambridge Neuropsychological Test Automated Battery (CANTAB Eclipse v6; Cambridge Cognition Ltd). The test battery comprised the Motor Screening Task (MOT), used to familiarise participants with the CANTAB system; the Choice Reaction Time (CRT), a simple 2-choice RT measure encompassing uncertainty; the Stop Signal Task (SST), used to measure response inhibition; the Attention Switching Task (AST), used to assess attention and cognitive flexibility; and the Spatial Working Memory (SWM) task, which measures retention and manipulation of visuospatial information. Data from these tasks did not inform any preregistered hypotheses but were collected for exploratory analyses and to describe the sample characteristics.

Screening

All participants were administered the Autism Quotient (AQ-10; Allison, Auyeung & Baron-Cohen, 2012) to screen for suspected autism; where a cut-off score of seven was used to exclude participant data. In addition to the AQ-10, OA also completed a dementia screening using the Mini Addenbrooke's Cognitive Examination (M-ACE; Hsieh et al. 2014). Participants scoring 25 or less on the M-ACE were excluded from the final sample; this resulted in one OA being excluded from the final analysis.

Procedure

First, participants completed a short training session (see below) followed by two 9-minute blocks of the ToM task, with a self-paced break between each. After completing the second block, participants completed the IRI followed by a 15-minute enforced break. Next, EF was evaluated using the CANTAB. The order of CANTAB testing was as follows: MOT, SST, AST, SWM and CRT. The SST, AST and CRT required the use of a left/right response button box, while the MOT and SWM were completed using the CANTAB touch screen. After a further 15-minute enforced break, participants completed block 3 and 4 of the ToM task. Lastly, the AQ-10 was administered and then OA completed the M-ACE. The total duration of the session was approximately 3 hours, which allowed for numerous breaks. This time also permitted casual interaction and refreshments with participants to reduce fatigue and increase engagement (see Results for explicit tests showing no significant cross-group fatigue effects).

The practice session comprised a self-paced document that contained a mock-up of the task, including two self-paced and one timed trial. These three practice trials included written feedback regarding what had just been shown and what the appropriate response should have been. After completing this, all further interactions with the ToM task were administered via E-Prime. The task started with a practice block consisting of 10 trials which followed the precise timing of the full experiment. In these practice trials only, participants received auditory feedback indicating the accuracy of their responses. The ToM training and ToM task proper were administered on a 17-inch Lenovo monitor powered by a Windows i7 machine. Participants responded using a standard 102 PC keyboard.

Statistical Analysis

All confirmatory analyses were conducted in SPSS (version 24) and JASP (version 0.12.2). Our primary hypotheses were assessed by running a 4-way mixed ANOVA on the ToM task data. This comprised Age Group as a between-subjects factor and three within-subjects factors: KoR, OOC and SOC. Table 3.2 below outlines the preregistered hypotheses and statistical tests used to assess these.

Table 3.2.
Summaries of preregistered hypotheses, predictions, and associated tests.

Hypothesis No.	Hypothesis	Predicted Direction of Costs[a]	Test
RH1	Explicit Knowledge of Reality will cause interference	$KoR_1 > KoR_0$	ANOVA main effect of KOR
RH2	Incongruent Self and Other knowledge states will be effortful	$SOC_1 > SOC_0$	ANOVA main effect of SOC
RH3	Knowledge of Reality will interfere when Self-Other perspectives are incongruent	SOC_1: $KoR_1 > KoR_0$	ANOVA interaction KoR*SOC
RH4	Managing an incongruent Self-Other perspective (a FB; SOC_1) will be more effortful than managing competing, alternate cued locations (OOC_1)	SOC_1 (FB) > OOC_1	t-test of two specific conditions (Fig. 3.3)
RH5	Ageing will be associated with an overall reduction in performance	Older > Younger	ANOVA main effect of Age Group
RH6	Ageing will be associated with reduced performance in aspects of conflict in ToM	KoR_1: Older > Younger SOC_1: Older > Younger OOC_1: Older > Younger	ANOVA KoR * Age Group ANOVA SOC * Age Group ANOVA OOC * Age Group

Note: Table presents a summary of statements taken from the original preregistration. KoR = Knowledge of Reality (0 = unknown; 1 = known). SOC = Self-Other Congruence (0 = congruent; 1 = incongruent). OOC = Other-Other Congruence (0 = congruent; 1 = incongruent). [a]Processing costs are inferred on the basis of increased reaction-times and error-rates. RH = research hypothesis.

Results

Sample Characteristics

OAs showed poorer performance across all neuropsychological measures; however, there was no statistically significant difference in self-reported perspective taking (PT) in the IRI; see Table 3.3. Due to equipment failure, no CRT data were acquired for one OA.

Table 3.3.

Sample characteristics by group: perspective taking and EF

	Age Group				
	YA		OA		
	M	*SD*	*M*	*SD*	*t*
IRI PT	18.66	4.49	18.40	4.79	.280
SST	162.23	46.75	196.74	38.27	4.039***
AST	42.43	35.97	69.56	58.05	2.809**
SWM	30.10	6.49	35.76	4.26	5.155***
CRT	303.89	52.02	369.36	52.55	6.229***

Note. YA = Young Adults. OA = Older Adults. PT = Perspective Taking, self-report measure from the Interpersonal Reactivity Index (IRI), max score 28 (higher score = higher perspective-taking proficiency). CANTAB measures: SST = Stop Signal Task, stop signal reaction time in ms. AST = Attention Switching Task, mean congruency cost; higher values indicate greater difficulty with managing attentional conflict. SWM = Spatial Working Memory, strategy score; higher scores represent poorer strategic performance. CRT = Choice Reaction Time, mean motor-response latency for correct responses in ms. Statistical significance based on independent t-tests comparing Age Group on each measure *p<.05. **p<.01. ***p<.001.

ToM Task Analyses

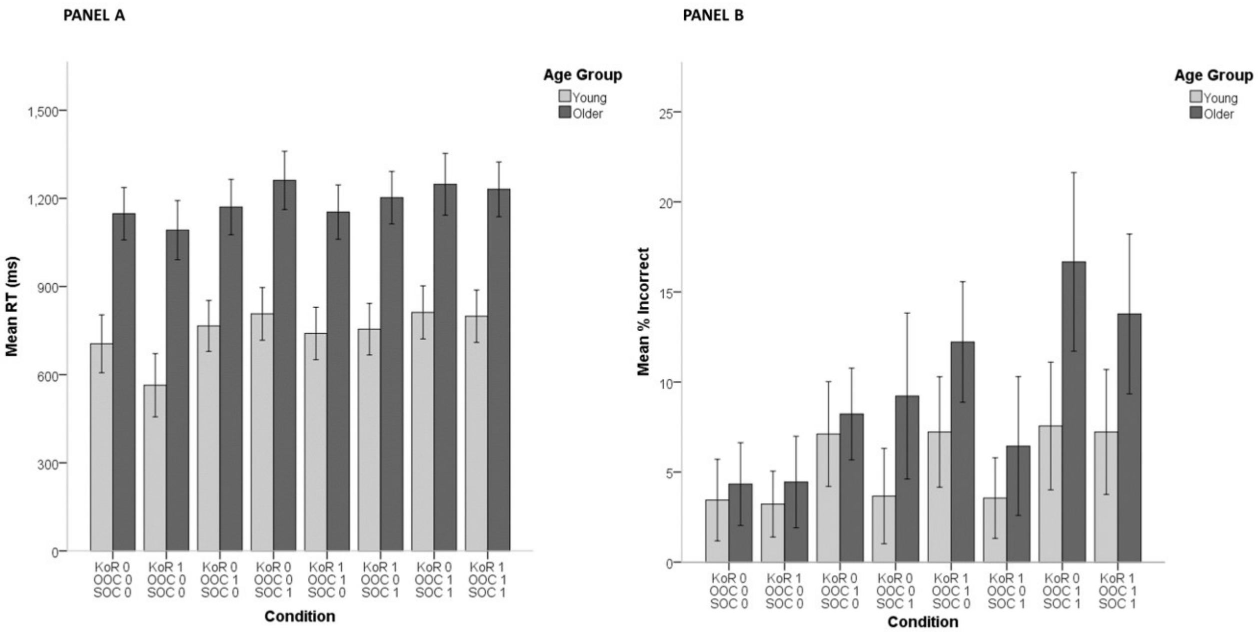
False Belief Task Data Pre-Processing

Prior to statistical analysis, the data were pre-processed as described in the study preregistration. No participants scored below chance in the 'filler' trials (< 31 based on a binomial probability distribution, $p < .05$), indicating that all participants were attending to the task and could therefore be included in the subsequent analyses. Next, only the ToM trials (not the 'filler' trials), where a correct response was given, were analysed. Trials with a response latency of ≤ 5ms were removed, which resulted in two trials being excluded (both OAs from the $KoR_1OOC_0SOC_0$ condition). Finally, RTs that were beyond 2 SDs from each participant's condition mean were removed (322 for YAs, 309 for OAs; 631 in total). Then, trials with incorrect responses – including null responses – were removed (387 for YA; 678 for OAs; total 1065 trials, 7.4% of overall dataset) from the RT analysis and analysed separately, as the amount of errors per condition. Altogether, 1698 trials (11.8%) were removed prior to analysis; 709 for YAs and 989 for OAs.

False Belief Task RT and Error-Rate Analyses

Our primary hypotheses were tested using a series of factorial analyses conducted on the RT and accuracy data. The condition-mean RTs (Figure 3.2A) were entered into a four-way mixed ANOVA with Age (young versus older) as a between-subjects factor, and three within-subjects factors: KoR ((reality unknown (KoR_0) versus known (KoR_1)); OOC ((agents' beliefs congruent (OOC_0) versus incongruent (OOC_1)); SOC ((agent's-participant's beliefs congruent (SOC_0) versus incongruent (SOC_1)). Similarly, a four-way mixed ANOVA was conducted on the error-rate data (Figure 3.2.B). To be consistent with our preregistered design and analysis protocol, we focus on the three repeated-measures main effects (KoR, OOC, SOC), the interaction between KoR and SOC, and the relationship between Age Group and our primary within-subjects manipulations.

However, for transparency, all results are reported in Table 3.4 below.



PANEL C

Condition Summaries

	Conflict	
Condition	low	high
Knowledge of Reality (KoR)	reality unknown (KoR$_0$)	reality known (KoR$_1$)
Other-Other Congruence (OOC)	congruent (OOC$_0$)	incongruent (OOC$_1$)
Self-Other Congruence (SOC)	congruent (SOC$_0$)	incongruent (SOC$_1$)

Figure 3.2. Panel A: Mean condition latencies separated by Group. Panel B: Mean percentage incorrect (accuracy) separated by Group. Error bars: +/- 2 SE. Panel C: Summaries of experimental each factor and level. Dark grey bars represent older adults and light grey represent younger adults.

Table 4.
Mixed ANOVA results for reaction time and error rate

	ANOVA: Reaction Times (milliseconds)					Sig. Post-hoc Comparisons RT		ANOVA: % Inaccurate Responses					Sig. Post-hoc Comparisons Inaccuracy	
Effect	df	df resid	F	Sig.	Partial ηp2	Sig. Effect(s) Direction[a]	Sig. Effect(s) Δ[a]	df	residual	F	sig.	Partial ηp2	Sig. Effect(s) Direction[a]	Sig. Effect(s) Δ[a]
Between-Subjects Effects														
Main effect of Age Group[RH6]	1	98	48.71	<.001***	.332	Young < Older***	-444.94	1	98	4.61	.034*	.045	Young < Older*	-4.04
Within-Subjects Effects														
Main effect of KoR[RH1]	1	98	47.09	<.001***	.325	$KoR_0 > KoR_1$***	47.59	1	98	.39	.536	.004	N/A	-
Main effect of SOC[RH2]	1	98	117.10	<.001***	.544	$SOC_0 < SOC_1$***	-96.96	1	98	7.58	.007**	.072	$SOC_0 < SOC_1$**	-2.24
Main effect of OOC	1	98	33.75	<.001***	.256	$OOC_0 < OOC_1$***	-48.22	1	98	61.45	<.001***	.385	$OOC_0 < OOC_1$***	-5.21
Interaction Effects														
KoR * Age Group[RH6]	1	98	2.25	.137	.022	N/A	-	1	98	.09	.769	.001	N/A	-
SOC * Age Group[RH6]	1	98	.062	.803	.001	N/A	-	1	98	5.98	.016*	.058	Older: $SOC_0 < SOC_1$***	-4.33
													SOC_1: Young < Older*	-6.03
OOC * Age Group[RH6]	1	98	7.94	.006**	.075	OOC_0: Young < Older***	-468.32	1	98	4.46	.037*	.044	OOC_1: Young < Older*	-5.44
						OOC_1: Young < Older***	-421.55						Young: $OOC_0 < OOC_1$***	-3.81
						Young: $OOC_0 < OOC_1$***	-71.61						Older: $OOC_0 < OOC_1$***	-6.61
						Older: $OOC_0 < OOC_1$*	-24.83							
KoR * SOC[RH3]	1	98	3.74	.056	.037	N/A	-	1	98	7.80	.006**	.074	KOR_0: $SOC_0 < SOC_1$***	-3.50
													SOC_1: $KoR_0 > KoR_1$*	1.53
KoR * OOC	1	98	36.05	<.001***	.269	KOR_0: $OOC_0 < OOC_1$*	-18.83	1	98	2.10	.150	.021	N/A	-
						KOR_1: $OOC_0 < OOC_1$***	-77.61							
						OOC_0: $KOR_0 > KOR_1$***	76.99							
						OOC_1: $KOR_0 > KOR_1$*	18.20							
SOC * OOC	1	98	20.07	<.001***	.170	SOC_0: $OOC_0 < OOC_1$***	-80.29	1	98	1.23	.271	.012	N/A	-
						OOC_0: $SOC_0 < SOC_1$***	-129.03							
						OOC_1: $SOC_0 < SOC_1$***	-64.89							
SOC * OOC * Age Group	1	98	4.32	.040*	.042	Young OOC_0: $SOC_0 < SOC_1$***	-146.14	1	98	1.41	.237	.014	N/A	-
						Young OOC_1: $SOC_0 < SOC_1$***	-52.25							
						Young SOC_0: $OOC_0 < OOC_1$***	-118.56							
						Young SOC_1: $OOC_0 < OOC_1$***	-24.66							
						See Table 5								
KoR * SOC * Age Group	1	98	4.19	.043*	.041	Young KOR_0: $SOC_0 < SOC_1$***	-73.87	1	98	6.81	.010**	.065	Older KOR_0: $SOC_0 < SOC_1$***	-6.67
						Young KOR_1: $SOC_0 < SOC_1$***	-124.52						Older SOC_0: $KOR_0 < KOR_1$*	-2.06
						Young SOC_0: $KOR_0 > KOR_1$***	83.33						Older SOC_1: $KOR_0 > KOR_1$*	2.83
						Young SOC_1: $KOR_0 > KOR_1$***	32.67						see Figure 4C/D	
						See Figure 4A/B								
KoR * OOC * Age Group	1	98	3.62	.060	.036	N/A		1	98	1.87	.175	.019	N/A	-
KoR * OOC * SOC	1	98	1.83	.180	.018	N/A	-	1	98	2.36	.127	.024	N/A	-
KoR * OOC * SOC * Age Group	1	98	2.10	.150	.021	N/A	-	1	98	1.35	.248	.014	N/A	-

Note. Table presents results from two separate mixed ANOVAs (RT and error-rate) and subsequent comparisons required to test statistically significant effects. [a]Sig. Effect(s) direction is only stated for simple main effects where the pairwise difference is statistically significant at $p_{bonf} < .05$. Δ = mean difference. Statistically significant 3-way interactions were probed using further repeated-measures ANOVAs as outlined in text. RH: Specific hypotheses were preregistered for these results, where each analysis (RH1, RH2 etc) is discussed further within the formal study preregistration; RHn refers to the research question and hypothesis number listed within the preregistration protocol. KoR = Knowledge of Reality (0 = unknown; 1 = known). SOC = Self-Other Congruence (0 = congruent; 1 = incongruent). Other-Other Congruence (0 = congruent; 1 = incongruent). N/A = non-applicable. *p<.05. **p≤.01. ***p≤.001.

Which factors are responsible for variation in the difficulty of reasoning about the beliefs of others generally?

Knowledge of Reality

RH1: Based on the theory that own knowledge can cause interference – a Curse of Knowledge (CofK; Birch & Bloom, 2004; 2007) – we predicted longer latencies and more errors when reality was known versus unknown. To assess this, we tested the within-subjects (w/s) main effect of KoR (prediction: $KoR_1 > KoR_0$). As detailed in Table 3.4, the effect of KoR was statistically significant; however, contrary to our preregistered hypothesis, participants were slower when they did not know reality (effect: $KoR_1 < KoR_0$), although there was no statistically significant effect of KoR to accuracy.

Incongruent Self-Other Perspectives

RH2: Prior research suggests that, when Self-Other perspectives differ, own perspective may interfere, and would thus need to be inhibited (Hartwright et al., 2015; Samson et al., 2005; Samson et al., 2015). On this basis, we predicted that ToM reasoning would be more effortful when Self and Other knowledge-states were incongruent (prediction: $SOC_1 > SOC_0$; RH2). To assess this, we tested the w/s effect of SOC (prediction: $SOC_1 > SOC_0$). In line with our preregistered hypothesis, participants were slower and made more errors when Self-Other perspectives where incongruent (SOC_1) versus congruent (SOC_0); Table 3.4.

Salient, Conflicting Knowledge

RH3. Prior research shows that FB reasoning is more effortful than TB reasoning (Apperly et al., 2008; Apperly et al., 2011); though the basis of this competition is unclear: is it due to having a privileged knowledge of reality – a CofK – or the mismatch between Self-Other perspectives? We tested the prediction that own knowledge of reality would interfere when Self-Other perspectives differed, by assessing the two way interaction between SOC and KOR. We predicted that a main effect of SOC would be qualified by an interaction with KOR, where error-rates and response latencies would be increased when Self-Other knowledge states were incongruent and reality known (prediction: SOC_1: $KoR_1 > KoR_0$).

Contrary to the CofK theory, the expected two-way interaction was not supported in RT. There was, however, a significant interaction in accuracy between KoR and SOC (Table 3.4): when reality was unknown, participants were more error-prone representing an incongruent versus congruent belief (effect KOR_0: $SOC_0 < SOC_1$), and when Self-Other knowledge states were incongruent, errors increased when reality was unknown versus known (effect SOC_1: $KOR_0 > KOR_1$). These two-way interactions were qualified by a three-way interaction between KOR, SOC and Age Group in RT and accuracy, which does lend support to the CofK, albeit in a more complex way.

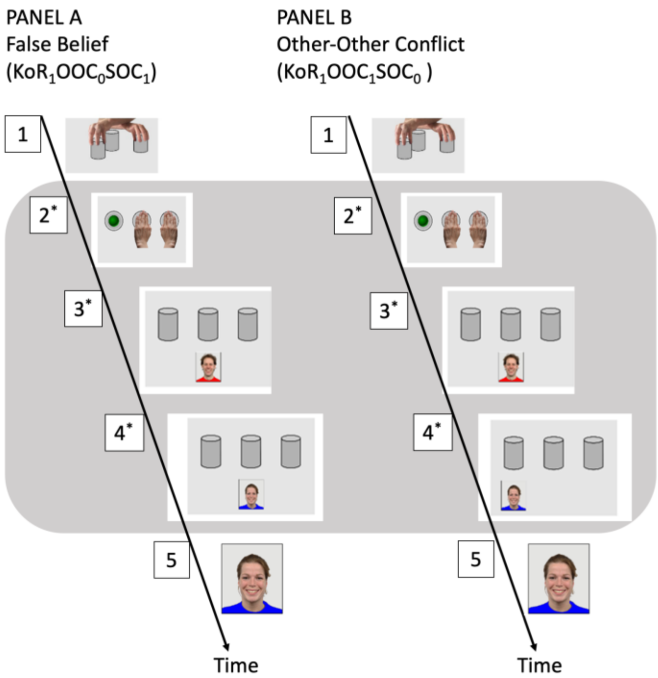
How does attentional cueing contribute to performance costs on false belief tasks?

RH4. Responding correctly in a FB paradigm typically requires a participant to shift attention between competing locations, while remembering which outcome maps onto which location. This effortful shifting between competing cued locations is a typical feature of FB paradigms though, unlike incongruence between Self- and Other-perspectives, it is not an essential feature of FB problems. We therefore developed a condition which required the participant to keep in mind, and shift between, two locations, without the need to represent the target agent's FB (OOC_1). Based on the Selection-Processing theory of FB reasoning (Friedman & Leslie, 2005), and work showing that incongruent Self-Other Perspectives create conflict distinct from other sources of conflict within ToM tasks (Hartwright et al., 2015; Samson et al., 2005; Samson et al., 2015), we hypothesised that there would be greater cognitive effort associated with holding in mind a competing Self-Other perspective (a FB, SOC_1), versus managing alternate cued locations (OOC_1). To assess this hypothesis, we compared performance in these two specific conditions within our ToM task (see Figure 4.3 below). In line with our predictions, a paired t-test revealed that a greater RT cost, around 32 ms, was observed when representing a FB ($KoR_1OOC_0SOC_1$; $\bar{X}$ = 978.53 ms) – compared to conflict from managing alternate locations ($KoR_1OOC_1SOC_0$; $\bar{X}$ = 946.88 ms), $t(99) = 2.61$, p = .010, d = .083. However, almost 5% more errors were made when managing alternate locations ($KoR_1OOC_1SOC_0$; $\bar{X}$ = 9.72%) versus managing a FB ($KoR_1OOC_0SOC_1$; $\bar{X}$ = 5.00%), $t(99) = 4.23$, p < .001, d = .415.

Which factors are responsible for variation in the difficulty of reasoning about the beliefs in healthy ageing?

Age-Related Differences in Response Speed & Accuracy

RH5. We predicted that OA would generally have longer RTs and increased error rate than YA, which we tested via the between-subjects main effect of Age Group. This hypothesis was supported, with OA generally providing slower, less accurate responses than YA (Table 3.4).



PANEL C

Event sequence ID (see Panels A and B)	Cup location cued: False Belief	Cup location cued: Other-Other Conflict
2	1	1
3	2	2
4	2	1
Total unique locations cued	2	2

Figure 3.3. Panel A: Schematic illustration of the key events in a false belief trial ($KoR_1OOC_0SOC_1$) and a conflicting Other-Other perspectives trial ($KoR_1OOC_1SOC_0$; Panel B). In the classic FB condition (SOC_1; panel A), two locations are cued – where the ball really is, and where the agent falsely believes the ball is – creating incongruence in Self-Other perspectives. In the alternative condition (OOC_1; panel B), two competing locations are cued by two agents with competing perspectives but, unlike the original FB, the target agent's and participant's beliefs are congruent. Each trial exemplar in Panels A and B is indicative of the events for each of the two trial types. A typical event sequence might proceed as: 1. Cups shuffled. 2. Reality status of one cup is indicated. 3. Red agent indicates their belief status regarding one cup. 4. Blue agent indicates their belief status regarding one cup. 5. Target for participant representation is indicated as the blue agent. *Note that the order of presentation of event IDs 2, 3 and 4 was counterbalanced and randomized across the full experiment. The grey shaded area highlights the critical components of a trial that are manipulated to generate either a false belief (panel A) or a conflicting Other-Other perspective (panel B). Panel C highlights which locations are cued in which aspect of the event sequence to illustrate the number of unique locations cued. . KoR = Knowledge of Reality (0 = unknown; 1 = known). SOC = Self-Other Congruence (0 = congruent; 1 = incongruent). OOC = Other-Other Congruence (0 = congruent; 1 = incongruent).

How do different sources of conflict differentially affect OA performance in FB tasks?

RH6. We predicted that OA would show greater difficulty than YA with self-perspective inhibition, handling incongruence between beliefs and managing conflict from multiple cued locations. These predictions were tested by assessing the two-way interactions between Age Group and each of the w/s factors: KoR, SOC and OOC. Contrary to our expectations; however, there were no statistically significant two-way interactions between KoR and Age Group in RT or error-rate, or between SOC and Age Group in RT. Age did, nonetheless, interact with SOC in accuracy, and OOC in both RT and accuracy. As detailed in Table 3.4, simple effects analyses demonstrated that accuracy was affected in both groups by manipulating Other-Other perspectives. Notably, the effect was almost doubled in OA, suggesting greater cognitive burden of OOC on OA ($OOC_0/_1$: $\bar{X}$ difference YA = 3.81%; OA = 6.61%). Indeed, OA made 5.44% more errors than YA when responding to a conflicting OOC. Further, only OA showed an effect of Self-Other perspectives on accuracy (OA: $SOC_0 < SOC_1$; $\bar{X}$ difference = 4.33%), suggesting the main effect of SOC on accuracy is driven by OA performance. Just as with OOC, OA made more errors than YA when responding to incongruent perspectives (SOC_1: YA < OA $\bar{X}$ difference YA = 6.03%). Regarding RT, both age groups slowed significantly to resolve a competing versus congruent Other-Other perspective (YA/OA: $OOC_0 < OOC_1$); although, contrary to accuracy, the effect of OOC on RT was more marked in YA than OA ($OOC_{0/1}$ $\bar{X}$ difference YA = 71.61ms; OA = 24.83ms).

There were several interaction effects with Age Group which we had not predicted and should therefore be considered exploratory. As detailed in Table 3.4, there was a significant three-way interaction between Age, OOC and SOC in RT. Two separate repeated-measures ANOVAs to evaluate this indicated that the interaction between OOC and SOC was statistically significant in YA ($F(1, 49) = 41.38$, $p < .001$, η p2= .458) but not OA ($F(1, 49) = 1.95$, $p = .169$). In both groups, conflicting Self-Other perspectives ($SOC_1 > SOC_0$) and Other-Other perspectives ($OOC_1 > OOC_0$) were completed slowest. However, pairwise comparisons indicated that YAs RTs were more affected by SOC when two agents' perspectives were congruent versus incongruent, and by the congruency of Other-Other perspectives when Self-Other perspectives were aligned versus incongruent; see Table 3.5.

Table 3.5.

*Pairwise comparisons of RT in YAs for OOC * SOC*

	$\bar{X}_0$	$\bar{X}_1$	*X* difference (0-1)	*SE*	P_{bonf}
	SOC				
OOC_0	634.511	780.653	-146.142	13.104	< .001***
OOC_1	753.066	805.317	-52.251	7.749	< .001***
	OOC				
SOC_0	634.511	753.066	-118.550	14.802	< .001***
SOC_1	780.653	805.317	-24.664	6.807	0.001***

Note. *p<.05. **p≤.01. ***p≤.001. SOC = Self-Other Congruence (0 = congruent; 1 = incongruent). OOC = Other-Other Congruence (0 = congruent; 1 = incongruent).

There was also a three-way interaction between KOR, SOC and Age Group in RT and accuracy. These interactions were interrogated using a further four, two-way repeated-measures ANOVAs and pairwise comparisons (Figure 3.4); one for each age group separated by RT and accuracy. The interaction between KoR on SOC in YAs was statistically significant in RT ($F(1, 49) = 15.45$, $p < .001$, $\eta p2 = .240$) but not in YAs accuracy ($F(1, 49) = .04$, $p = .836$, $\eta p2 = .001$). Conversely, the interaction between KoR and SOC was non-significant in OAs RTs ($F(1, 49) = .00$, $p = .949$, $\eta p2 = .000$), but significant in OAs accuracy ($F(1, 49) = 9.07$, $p = .004$, $\eta p2 = .156$). The interaction effect in YA adults' RT was due to a more marked slowing when managing a CofK ($SOC_{0/1}$ $\bar{X}$ difference $KoR_0 = 73.87$ms; $KoR_1 = 124.52$ms; Figure 3.4A). Interestingly, both age groups slowed to manage the CoFK (KOR_1: $SOC_{0/1}$; Figure 3.4A/B), and both maintain within-group accuracy levels between TB and FB when reality was known (KoR_1: SOC_0 ≈

SOC_1; Figure 4C/D). However, specific to OA, a substantial cost to accuracy was associated with representing an incongruent- versus congruent belief when reality was unknown (KoR_0: $SOC_{0/1}$ $\bar{X}$ difference = 6.67%; Figure 3.4D), suggesting the two-way interaction between SOC and Age is driven by OAs poor performance in the reality unknown FB condition (KoR_0SOC_1). Indeed, both age groups take longest to resolve this reality unknown FB, indicating substantial cognitive demand.

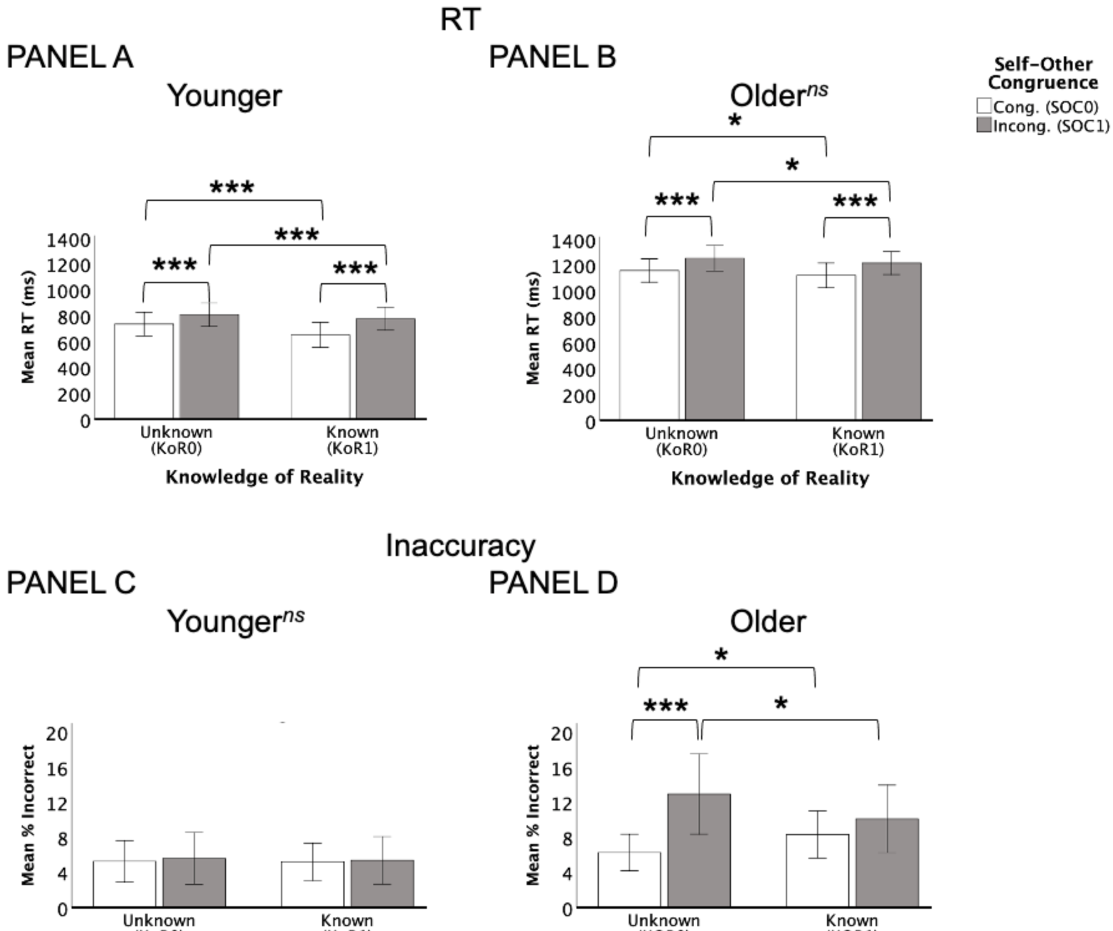


Figure 3.4. Estimated marginal means for RT and accuracy across YA and OA. Statistically significant pairwise comparisons across SOC and KOR indicated, where *p<.05. **p<.01. ***p≤.001. ^{ns}Interaction effect non-significant. Error bars: +/- 2 SE

Additional Exploratory Analyses

Knowledge of Reality, FB Reasoning and the 'Curse of Knowledge' (CofK)

The three-way interaction between KOR, SOC and Age Group suggested that our initial interpretation of a CofK should be revised. When knowing the ball's location, participants were slower to respond if the agent held a false-, (SOC_1) rather than a true-belief (SOC_0), which is consistent with a CofK. However, as shown in Figure 3.4, participants were slowest and most error-prone overall when reasoning about an agent with a FB when reality was unknown (KOR_0). Participants experienced greatest difficulty when representing an agent who falsely believed the ball was at a location it was clearly not, which seems counterintuitive to the CofK hypothesis. To further explore this, we theorised that belief representation in the reality unknown condition (KoR_0) could pose additional difficulty because the empty location – the cup labelled with an X – should be avoided, which would require additional selection and control processes. Leslie and colleagues proposed that, in the classic object transfer FB task, it is implicit that the target agent wants to find the object (Leslie et al., 2005). With the present task, regardless that the target agent had no awareness of the contents of any of the three locations – as in a typical FB task – here too it was implicit that the target agent would want to *avoid* the empty location. Knowing that the ball is not in a particular location bestowed the participant with privileged knowledge of where the ball definitely was *not*. With the current paradigm, we therefore effectively had two false belief conditions: the classic FB, as seen in the original object transfer task ($KoR_1OOC_0SOC_1$), and a novel, reality unknown FB, because the target agent thought that the ball was somewhere the participant knew for certain it was not ($KoR_0OOC_0SOC_1$). Should the participant need to inhibit their knowledge of the location to be avoided, it would be reasonable to expect greater cognitive costs associated with the novel, reality unknown FB ($KoR_0OOC_0SOC_1$)

versus the classic FB ($KoR_1OOC_0SOC_1$) condition. To test this, two further paired t-tests were performed, taking the data from all participants. Our assertion was supported in RTs (classic FB $\bar{X}$ = 978.53 ms; novel FB $\bar{X}$ = 1033.99 ms; $\bar{X}$ difference = 55.46 ms), $t(99) = 5.20$, $p < .001$, $d = .141$ and, though not statistically significant in accuracy, the direction was consistent with the RT data, ruling out a speed-accuracy trade-off (classic FB $\bar{X}$ = 5.00% errors; novel FB $\bar{X}$ = 6.45% errors; $\bar{X}$ difference = 1.45%), $t(99) = 1.83$, p = .070. We propose that this pattern could be indicative of a cognitively effortful 'double inhibition' (Leslie, German and Pollizi, 2005).

ToM, Ageing, and EF

Correlation analyses (for correlation matrix of all FB task and EF measures, see Fig. 3.7) suggested that RTs in all conditions were significantly positively correlated with individual differences in two-choice motor-response time (CANTAB CRT; r=.619 to .682), inhibitory control (CANTAB SST; r=.354 to .395) and spatial working memory (CANTAB SWM; r=.325-.417). No RTs were significantly correlated with attentional capacity (CANTAB AST; r=.113 to .168) or self-reported ToM (IRI PT; r=-.067 to -.148). For error-rate, only motor-response time was significantly correlated with all conditions (CANTAB CRT; r=.243 to .434)..

To assess which aspects of EF explain the magnitude of conflict introduced within each experimental factor, we derived a 'cost-factor' (cf) for each of the three factors, KOR, OOC, and SOC. We collapsed across task conditions (KoR/SOC/OOC) and subtracted those conditions within each factor with pre-supposed high levels of conflict ($KoR_1/OOC_1/SOC_1$) from those with lower levels of conflict ($KoR_0/OOC_0/SOC_0$):

$KoR_{cf} = (KoR_1OOC_0SOC_0 + KoR_1OOC_1SOC_0 + KoR_1OOC_0SOC_1 + KoR_1OOC_1SOC_1) – (KoR_0OOC_0SOC_0 + KoR_0OOC_1SOC_0 + KoR_0OOC_0SOC_1 + KoR_0OOC_1SOC_1)$

$OOC_{cf} = (KoR_0OOC_1SOC_0 + KoR_1OOC_1SOC_0 + KoR_0OOC_1SOC_1 + KoR_1OOC_1SOC_1) – (KoR_0OOC_0SOC_0 + KoR_1OOC_0SOC_0 + KoR_0OOC_0SOC_1 + KoR_1OOC_0SOC_1)$

$SOC_{cf} = (KoR_0OOC_0SOC_1 + KoR_1OOC_0SOC_1 + KoR_0OOC_1SOC_1 + KoR_1OOC_1SOC_1) – (KoR_0OOC_0SOC_0 + KoR_1OOC_0SOC_0 + KoR_0OOC_1SOC_0 + KoR_1OOC_1SOC_0)$

This was done separately for RT and for accuracy giving 6 CFs per participant. We ran six separate stepwise multiple regression analyses to predict each of the CFs. The data met the assumptions for multicollinearity, homoscedasticity and linearity, independence of errors (based on Durbin Watson ~2) and that the error terms were normally distributed. The OAs demonstrated more extreme CF values; however, all data were included: all participants had passed screenings for dementia and autism, all included ToM data had met performance criteria specified in the pre-registration, and the CF measures reflect a summary of an individual's repeated, consistent behavioural performance over numerous trials. Consequently, we considered all CF values were representative of typical task performance within a continuum of variability. Each regression analysis included six predictors: Age Group, the self-reported perspective taking (PT) measure from the IRI, plus the four neuropsychological measures from the CANTAB: SST, AST, SWM, and CRT (see Table 3.6). All six predictors were entered into the step-wise regression analysis and the models which explained the most variance with the fewest statistically significant variables—i.e. the most parsimonious models—were retained.

Table 3.6.
Multiple regression results for reaction time and error rate cost factors

	CF RT data										CF error-rate									
Cost Factor	df	df resid	F	Sig.	R^2	adj R^2	Sig. predictors	Std beta coef	t	Sig.	df	df resid	F	Sig.	R^2	adj R^2	Sig. predictors	Std beta coef	t	Sig.
KOR	-						*ns*										*ns*			
OOC	1	97	8.453	.005**	.080	.071	Age Group	-.283	-2.907	.005**	1	96	6.554	.002**	.120	.102	SWM	.264	2.761	.007**
																	IRI PT	-.231	-2.413	.018*
SOC	1	97	5.217	.025*	.051	.041	AST	.226	2.284	.025*	1	96	9.023	<.001***	.158	.141	CRT	.308	3.276	.001***
																	IRI PT	-.222	-2.363	.020*

Note: Table presents results from most predictive model, based on R^2, from six separate stepwise multiple regression analyses (RT and error-rate for each CF). Age Group was coded as a dichotomous categorical variable. PT = Perspective Taking, self-report measure from the Interpersonal Reactivity Index (IRI). CANTAB measures: SST = Stop Signal Task, stop signal reaction time in ms. AST = Attention Switching Task, mean congruency cost; higher values indicate greater difficulty with managing attentional conflict. SWM = Spatial Working Memory, strategy score; higher scores represent poorer strategic performance. CRT = Choice Reaction Time, mean latency for correct responses in ms. KoR = Knowledge of Reality (0 = unknown; 1 = known). SOC = Self-Other Congruence (0 = congruent; 1 = incongruent). OOC = Other-Other Congruence (0 = congruent; 1 = incongruent). ns - non-significant: no statistically significant predictive models identified. *p<.05. **p<.01. ***p<.001.

Table 3.6 shows that there were no significant predictors of KoR_{CF} in either RT or accuracy. Consistent with the earlier analyses, Age Group was identified as a significant predictor of OOC_{cf} RT, explaining 7% of the variation, where higher costs to RT were associated with lower age. Further, reduced self-reported ToM (IRI PT) and less efficient use of SWM (spatial working memory) explained around 10% of variation in OOC_{cf} error-rate. 4% of the increased cost introduced to SOC_{cf} RT was associated with greater difficulty managing attentional conflict (AST). Moreover, longer baseline RT (CRT) and poorer ToM proficiency could explain 14% of the increased error-rate associated with varying Self-Other perspectives.

Practice/fatigue effects

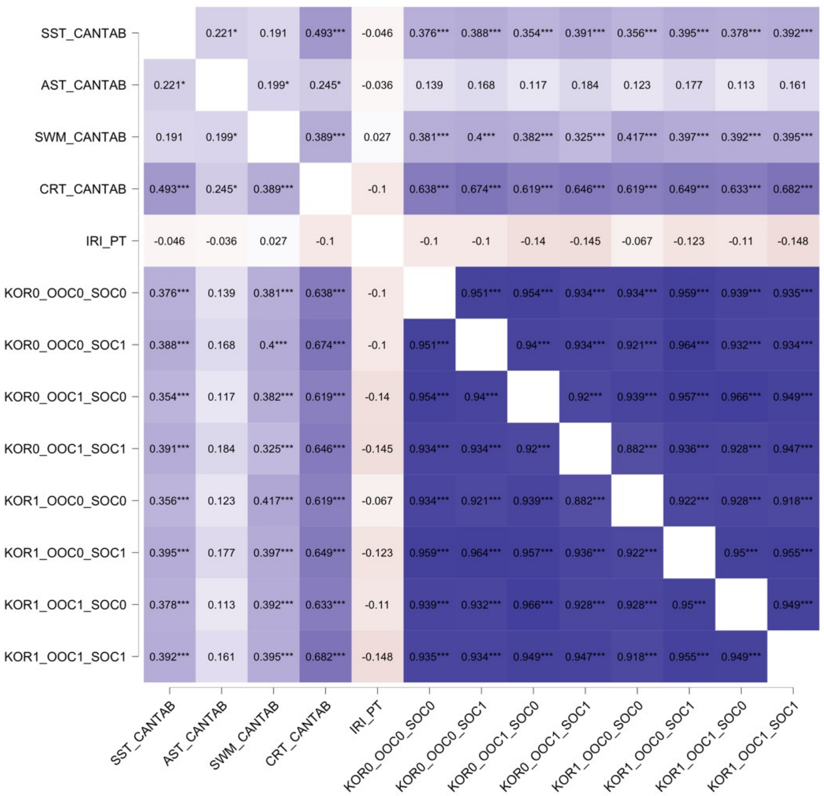
As the testing session was relatively lengthy (> 3 hours), we examined whether there was a disproportionate effect of learning or fatigue between the two age groups. A mixed ANOVA with task block (1, 2, 3, & 4) as a repeated-measures factor and Age Group (young vs older) as a between-subjects factor, revealed no significant interaction with either RT or error-rate (both *p*'s > .05) suggesting that there was no significant difference in any potential effects of learning or fatigue between the two groups.

Relationship between ToM performance, EF, and perspective-taking

Finally, correlational analyses were performed with RT, accuracy, EFs, and self-report perspective-taking. To this effect, correlation matrices are presented below, see Figure 3.5. Our core analyses were driven via our pre-registered analysis plan, which did not include examination of correlations between individual task conditions and EF measures;

thus these correlations are not interpreted further here, but are presented in Figure 3.5 for reference.

Panel A



Panel B

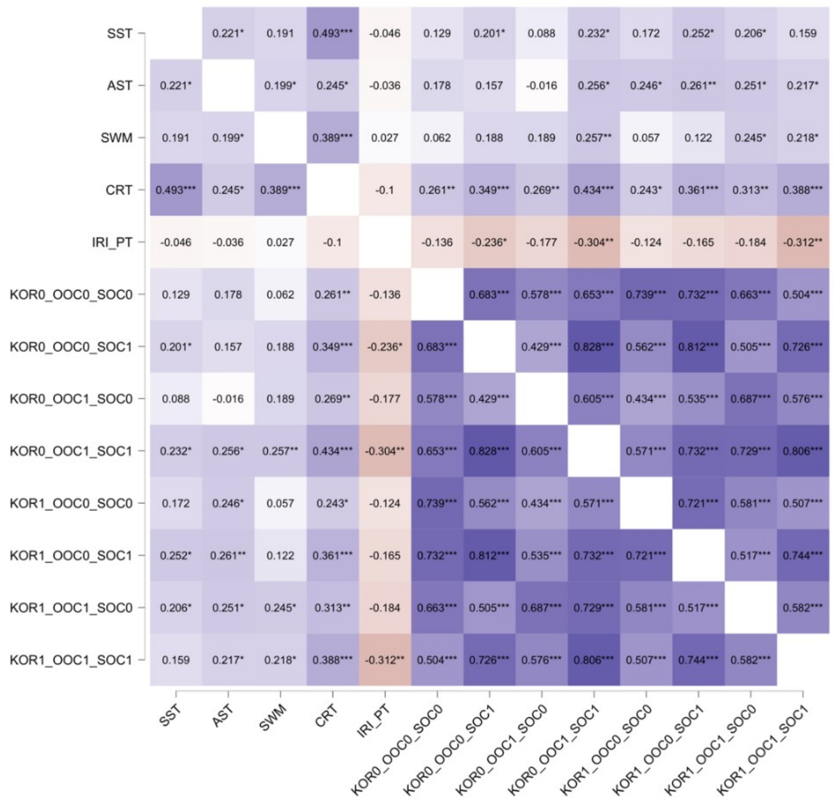


Figure 3.5. Relationship between RT (Panel A) and error-rate (Panel B) for each of the experimental conditions with EF and IRI PT; each value reflects correlation coefficients (rPearson's) across these. PT = Perspective Taking, self-report measure from the Interpersonal Reactivity Index (IRI), max score 28 (higher score = higher ToM proficiency). CANTAB measures: SST = Stop Signal Task, stop signal reaction time in ms. AST = Attention Switching Task, mean congruency cost; higher values indicate greater difficulty with managing attentional conflict. SWM = Spatial Working Memory, strategy score; higher scores represent poorer strategic performance. CRT = Choice Reaction Time, mean motor-response latency for correct responses in ms. KoR = Knowledge of Reality (0 = unknown; 1 = known). SOC = Self-Other Congruence (0 = congruent; 1 = incongruent). OOC = Other-Other Congruence (0 = congruent; 1 = incongruent). Figure produced using JASP v0.12.2., https://jasp-stats.org/ *p<.05. **p<.01. ***p<.001.

Discussion

Prior research has produced conflicting findings regarding whether ToM declines in healthy ageing (Henry et al., 2013; Love, 2015; Phillips et al., 2011). To unpick this, the present study assessed the role of three theoretically derived sources of conflict: privileged outcome knowledge (KoR), congruence of Self-Other perspectives (SOC), and competing cued locations (OOC). By assessing a series of preregistered hypotheses, this study highlights two important findings.

Competing Self-Other perspectives and competing cued locations tap different cognitive mechanisms, which affect YA and OA differently

We predicted that conflicting perspectives would be effortful (Apperly et al., 2008; 2011), particularly for OA (Bottiroli et al., 2016; German and Hehman, 2006), and that managing incongruent Self-Other perspectives would be more effortful than competing cued locations (Hartwright et al., 2015; Samson et al., 2005; 2015). Our data partially support these predictions, but the findings were more nuanced than expected. Unexpectedly, both groups showed similar slowing to resolve a competing Self-Other perspective, and both show comparable error-patterns when resolving the classic FB scenario. Conversely, though faster overall, YA slowed more than OA to manage invalid cueing, resulting in them committing substantially fewer errors than OA. This might first appear to reflect a speed-accuracy trade-off (SAT); however, as these behaviours differ through the type of conflict, the data are more consistent with meaningful processing differences between groups. Such SAT differences have been shown to have a neurological, rather than a strategic, basis (Forstmann et al., 2011). Indeed, further to divergent behavioural profiles, the cognitive systems co-opted to resolve each source of conflict were condition specific: managing competing Self-Other perspectives was supported by attentional systems, whereas invalid cueing drew from spatial working memory. Considering this pattern of differences, our work suggests that different mechanisms manage these two sources of conflict. This is consistent with neuroimaging data showing that representing a FB is functionally distinct from attentional demands due to cueing behaviourally relevant spatial locations (Mars et al., 2012; Scholz et al., 2009; Young, Dodell-Feder & Saxe, 2010). Our work uniquely shows, however, that because of these different mechanisms, the nature of conflict in ToM differentially affects speed-accuracy response behaviours across age groups. This can explain the appearance of poorer ToM in OA. Conflicting perspectives were resolved similarly, whereas OA were less reactive to invalid cueing, resulting in proportionately more mistakes. Our data indicates that individual differences in attentional capacity best explain RT performance when managing competing Self-Other perspectives, and that errors better reflect limitations in EF and motor-response speed, rather than an age-related decline in ToM proficiency per se. Our work highlights that OA were more susceptible than YA to irrelevant cues in a FB context, and that unnecessary demands on working memory, through the cueing of invalid locations, disproportionately affected OA performance. Given the association identified between age, cueing and working memory, limits on processing speed, which declines with age, may explain this pattern of behaviour (Brown et al., 2012; German & Hehman, 2006; Salthouse, 1996). Critically, therefore, prior reviews and meta-analyses of ToM performance in ageing should be carefully interpreted: studies where such cueing occurs may inflate age-related changes in ToM capacity, due to incidental task demands which disadvantage OAs.

Interference from Knowledge of Reality is not affected by age, but by the perceived higher order intentions of Other

We hypothesised that KoR could interfere with one's ability to reason, due to bias towards own self-perspective – a CofK (Birch & Bloom, 2004; 2007), and that OA would be disproportionately affected by KoR, as prior work suggests that the CofK is more pronounced in later life (Bernstein et al., 2011). However, these predictions were not realised: participants were *faster* when they knew where the ball was and both age groups performed comparably, regardless of KoR. To explore these unexpected findings, we considered studies of belief-desire reasoning, where the target agent may wish to avoid the given target object. Increased difficulty is associated with processing false- versus true-beliefs and avoid- versus approach-desires, where a FB combined with an avoidance desire attracts maximal processing costs (see Apperly et al., 2011; German and Hehman, 2006; Hartwright et al., 2012; Leslie et al., 2005). Mentalizing about an agent with a FB regarding an empty location would be doubly effortful, as participants must inhibit their knowledge that the agent must avoid the location they (falsely) believe to be true (Leslie et al., 2005). Our exploratory analysis supported this assertion: participants took longer to resolve a reality unknown FB compared with the classic FB, suggesting that KoR itself is not the cause of the CofK. Instead, participants' initial internal reference towards the agent's desire created conflict, in our case, resulting in re-direction away from the empty cup. This finding is consistent with work showing that we automatically anticipate others' behaviour will fulfil, rather than conflict, their desires (Ferguson and Breheny, 2011), which suggests that the CofK is mediated by a perception of the agent's higher order intentions.

Limitations and future directions

There are a number of limitations to consider for the current study. First, as we have seen from the meta-analysis in Chapter 2, education can potentially modify ToM ability in ageing. We previously showed that older adults who had greater levels of education (i.e., >14 years) fared significantly better than their less educated counterparts (<14 years). In the two empirical investigations presented in this thesis, greater than 50% the OA samples were educated to at least undergraduate degree-level, which suggests that the education level in our sample was greater than what is prevalent in the wider population within this demographic. That we did not statistically adjust for education is a limitation of our findings and future work would benefit from measuring and carefully controlling for the influence of education.

The current design used dichotomous groups comprising young and older adults, where young was defined as 18-29 years old and older was defined as 60-79. Previous ageing-ToM studies using a *lifespan* approach have used more constrained age categories and have included more than two groups—e.g., young, young-old, old-old (e.g., Laillier et al., 2019). The advantage of this approach is that it allows for a more fine-grained analysis of ToM modification across the ages, which subsequently leads to a greater understanding of what the trajectory of ToM capacities look like at multiple stages of the developmental journey, from childhood and adolescence to old age. In the current study, we were only able to compare young vs old, without being able to discern changes to social cognitive function in, for example, middle to middle-to-late adulthood. While such a design would require a large number of participants to reach statistical feasibility, future studies should look to adopt a lifespan approach to bring greater understanding of when ToM 'peaks' and precisely how it degrades through the years.

Conclusion

The present study suggests that FB reasoning is effortful for OA beyond the non-social cognitive demands of classic ToM investigations. Performance in each of our ToM scenarios paralleled individual differences in inhibitory control and spatial working memory. However, the magnitude of conflict experienced and the cognitive systems co-opted to resolve this were condition specific: managing competing cognitive perspectives was supported by attentional systems whereas invalid cueing appeared to draw on spatial working memory. Further, OA were particularly disadvantaged by invalid cueing. This indicates that prior studies may overestimate the effects of ageing on ToM and highlights the need for carefully managing conflict in future studies of ageing and ToM.

Chapter Four: Neural basis of interference in theory of mind reasoning in health ageing

Chapter précis and background

In the current chapter, we investigated the neurocognitive profile of ageing ToM by administering fMRI to both young (YA) and older adult (OA) participants while they responded to our novel false belief paradigm and an established ToM functional localizer. We adapted our false belief task - originally employed in our behavioural study (Chapter 3) – to only include two factors: Knowledge of Reality (KOR) and Self-Other Congruence (SOC). The KOR factor manipulated whether the participant had a privileged knowledge of reality (reality unknown vs. reality known) and the SOC manipulation systematically varied the congruence between Self and Other beliefs (congruent vs. incongruent). These factors were conceptually identical to Chapter 3. We investigated the effect of each of these mechanisms in the context of reasoning about a false belief, and examined whether the interference arising from these factors differentially affected older (compared to young) adults.

On a behavioural level, we found a main effect of KOR, where overall, participants were faster in resolving the false belief problem when reality was known (compared to reality unknown). This is a replication of our previous finding from Chapter 3. The KOR effect was driven by a KOR x Age interaction where the observed benefit of knowing reality was only present in OA. We did not find an effect of SOC, nor an interaction between SOC and KOR, or SOC and Age. From the neuroimaging data, we were able to show differentiation in the neural response to reality known (vs reality unknown), in a direction consistent with our behavioural findings. While we were unable to show an overall effect of SOC, we did observe an interaction in KOR x SOC where we examined participant brain responses to incongruence in self-other beliefs (vs congruence) when reality was unknown. The cortical regions highlighted within both of these tests generally fell within the well-established *mentalizing network*: primarily, bilateral temporoparietal junction, precuneus, left dorsolateral prefrontal cortex, and medial prefrontal cortex. We did not find an interaction with age based on either of the two above contrasts on a neural level, however, we were able to show emerging age-related differences in the recruitment of predefined cortical regions, as informed by a ToM functional localizer.

Overall, in this chapter, we showed knowledge of reality to be a potentially important factor in false belief tasks, where participants – and in particular, OA – showed poorer performance when they did not have a knowledge of reality. We also identified discrete nodes of the mentalizing network based on our theorised sources of interference in belief reasoning, and showed emerging between-group differences within this network, though these did not survive statistical thresholding.

Introduction

Central to human social functioning is the capacity to reason about the subjective mental states of others, which involves inferring others' goals, intentions, desires, and knowledge states (Premack & Woodruff, 1978). This ability – known as 'theory of mind' (ToM) – typically develops early on in the lifespan, usually around pre-school age and before the age of 4-5 in neurotypical children (see Scott & Baillargeon, 2017, for a review on early mental-state understanding). Beyond childhood, one's ability to mentalize or mindread others' mental states remains relatively intact until advanced age; where healthy, non-pathological ageing seems to have a deleterious effect on one's ToM (for reviews, see Henry et al., 2013; Moran, 2013).

There are, however, inconsistencies in the literature with respect to age-related decline in ToM. One of the earliest demonstrations of age-related changes in adult ToM was conducted by Happé, Winner, Brownell (1998), on two groups of participants, with mean age 21 and 73, respectively. Participants were given ToM (double-bluffs, faux pas, and white lies) and non-ToM stories to read and then answered questions based on their memories of the stories. Overall, while the older group was slower in responding to both ToM and non-ToM stories, older adults generally answered more questions correctly in the ToM condition compared to their younger counterparts. There was no difference in performance between the two groups in the non-ToM control condition. Happé and colleagues concluded that the increased accuracy observed in older adults is perhaps reflective of better or improved ToM abilities in advanced age.

In response to Happé and colleagues, Maylor et al. (2002) tested young and older participants with Happé et al's 'Strange Stories' design. Crucially, Maylor and colleagues matched the two age groups in terms of education and also systematically varied the need to rely on memory to answer the questions; in one condition, the story was presented alongside the questions such that the participants could refer back to the narrative, should they wish – eliminating the memory load present in Happé et al (1998). Across two experiments, Maylor et al. found that the young group outperformed the older group for ToM stories with a memory load. Performance between the two groups did not differ in a non-ToM, control condition. Overall, Maylor et al.'s findings were the first to experimentally evidence an age-related decline in ToM between young and older adults.

Subsequent and contemporary literature has generally converged on a consensus that ToM abilities seem to decay with healthy ageing (Bottiroli et al., 2016; Hughes et al., 2019; Phillips et al., 2011; Reiter et al., 2017). For example, German and Hehman (2006) tested young and older adults across two ToM experimental manipulations: false (FB) vs true belief (TB) conditions. Successfully resolving a FB problem requires an appreciation that social actions are sometimes based on beliefs and intentions (as opposed to *reality*), even when those beliefs are evidently false (Wellman, Cross, & Watson, 2001). A TB, on the other hand, can usually be resolved based on one's own true (and sometimes privileged) knowledge of reality, and is therefore typically operationalised as a non-social, non-ToM control condition in FB investigations. In light of this, German and Hehman (2006) found that, while there was a universal cost associated with resolving a false (versus true) belief, the cost was disproportionately larger in the older adult group.

More recently, Laillier and colleagues (2019) tested 60 healthy adults aged 20-75 using the MASC (Movie for the Assessment of Social Cognition; see Dziobek et al., 2006). The MASC is a 15-minute video-based ToM paradigm where participants watch a film about four characters at a dinner party. The film is paused 46 times and, on each pause, participants are asked about the characters' thoughts, intentions, and desires.

The use of a continuous and dynamic film-based paradigm is purported to have greater ecological validity, which better mimics real-life social interactions compared to static cartoon-based tasks (Dziobek et al., 2006). Accordingly, Laillier et al. reported an age-related decay of ToM across the adult lifespan, and while cognitive measures (processing speed and episodic memory) partially predicted mindreading ability, they did not mediate the overall effect of age on ToM competence.

The extent of the ToM age effect, however, is not uniform (see meta-analysis, Chapter 2). There is substantial heterogeneity in the magnitude of the age effect depending on features of the task design and/or demographic make-up of the OA group (e.g., education, age, sex ratio, etc.). Indeed, some studies have reported no age differences at all after controlling for variability in global cognition and/or executive function (EF; see Chapter 2). Further, the degree to which EFs are recruited differ greatly across tasks, and even within the same modality of tasks the demand on executive resources varies. Often, these demands are conflated or masked behind social manipulations of a task, and are therefore difficult to disentangle. There is also heterogeneity in the neuroimaging literature in terms of the activated networks during different ToM tasks, presumably due to the variability in EF and social demands. In their meta-analysis of the most common ToM tasks, Schurz et al. (2014) reported significant inconsistencies in the recruited networks, and found convergence across tasks only in the posterior TPJ and dorsal mPFC.

As such, from the extant literature, it is unclear as to what the precise neurocognitive mechanisms are that are responsible for variation in performance, in young versus older adults. What processes involved in reasoning about others' mental states pose greater difficulty for older adults?

Manipulating the (in)congruence between self and other beliefs, and adjusting the salience of one's own belief, has highlighted key neurocognitive mechanisms involved in the inference of mental states (e.g., see Hartwright, Apperly, & Hansen, 2014, 2015). Resolving a FB problem requires the disengagement or inhibition of a self-evident truth (or knowledge of reality) in order to represent the FB of the other. Conflict arising from competing self-other beliefs and a knowledge of the true state of affairs taps working memory, attention, and inhibition, to support mental state representation (Austin, Groppe, & Elsner, 2014; Carlson, Moses, & Breton, 2002; Leslie, Friedman, & German, 2004). Most pertinent in resolving FB problems is the processing resource of conflict inhibition or inhibitory control. In a typical FB task, to effectively attribute and represent the belief of a third party, the participant must hold in mind an erroneous belief while simultaneously disregarding salient knowledge of an evidently TB. The TB selection is the more salient candidate in this scenario – due to an egocentric bias, see below – but through the application of inhibitory control the salient TB is disregarded in favour of the FB (see Friedman & Leslie, 2004). Recent EEG work suggests that we may initially anchor to our own, informed knowledge of the world before representing others' perspectives; thus, we must overcome this initial egocentric bias to assume someone else's belief (Bradford, Brundson, & Ferguson, 2020). Indeed, a range of studies have found inhibitory control to be a predictor of FB performance – though, these findings have mostly come from paediatric samples (Flynn, O'Malley, & Wood, 2004; Mutter, Alcorn, & Welsh, 2006; Perner, Lang, & Kloo, 2002). When reality is known – i.e., when one has privileged knowledge of an event's outcome – ToM performance tends to be compromised due to greater egocentric bias, a common fallacy in belief reasoning known as the *curse of knowledge* (see Birch, 2005; Birch & Bloom, 2007) or *reality bias* (Mitchell et al., 1996). Inhibiting this salient self-perspective to represent the other's FB recruits additional processing, and there is evidence to suggest that ‘self-perspective inhibition’ can be broadly dissociated from wider executive functions (Hartwright et al., 2015; Samson, Apperly, Kathirgamanathan, & Humphreys, 2005; Samson, Houthuys, & Humphreys, 2015).

Conflict arising from inconsistent Self-Other beliefs can be further demonstrated in studies of perspective-taking. Ferguson, Apperly, and Cane (2017) administered a Level 1 perspective-taking task where participants were required to confirm the number of discs in a visual scene. Participants resolved the task through either their own (egocentric) or via an avatar’s (altercentric) perspective, where the two perspectives were either consistent or inconsistent. Ferguson and colleagues revealed differential performance based on consistency, with higher accuracy rates when participant and avatar shared the same perspective.

In our own work (Rahman et al., 2021; Chapter 3), we previously manipulated the salience of self-knowledge and congruence of self-other perspectives in studying the source of interference in FB reasoning, and examined how the nature of these interferences may interact with ageing. We found that resolving incongruence (vs. congruence) between self-other perspectives was universally difficult, however, the cost of incongruence was disproportionately larger for older adults. We further found that this overall cost – incongruence in self-other beliefs – was predicted by performance on an attention switching task.

However, the neural bases of the effect of an explicit knowledge of reality and incompatibility in self-other beliefs, and how these sources of interference in FB reasoning interact with age, is still not fully understood. More precisely, whether young and older participants show differential recruitment of neural regions in managing these sources of competition is largely underexplored. In the current experiment therefore, using a modified version of the original FB task, we first attempted to replicate the findings our behavioural study (Chapter 3). Second, we investigated differential neural responses to the sources of interference highlighted previously. Importantly, we examined whether these effects interacted with age and whether there were between-group differences in the activation of circumscribed regions of the brain, as informed by an extensively-published ToM functional localizer (Saxe & Kanwisher, 2003).

The ToM Network

While there is a general consensus on a theoretical definition of ToM (i.e., *imputing mental states to self and others*; Apperly, 2012), there lacks a specificity on what ToM paradigms *actually* measure (see Quesque & Rosetti, 2020). There is generally a large heterogeneity in the tasks employed in the pursuit of estimating ToM and the tasks themselves vary as to whether they require mental state representation. Indeed, some purported ToM tasks require no mental state representation at all (Quesque & Rosetti, 2020). Even when the imputation of others' mental and knowledge states *is* required, there is again variation in the modality of tasks employed (e.g., verbal or non-verbal, static vs. dynamic, etc.); with each measure potentially tapping separate and distinct executive resources. Despite this practical difficulty in pinpointing and operationalising ToM into a specific, homogenous concept, imaging research on the neural basis of ToM generally points towards a core 'mentalizing' network (see Table 4.1; Abu-Akel & Shamay-Tsoory, 2011; Schurz et al., 2014).

Table 4.1

Theory of mind brain regions (from Abu-Akel and Shamay-Tsoory, 2011; reproduced with permission)

Brain region	Brodmann's area
Posterior regions	
Temporo-parietal junction (incl. Inferior parietal lobe) (IPL/pSTS or TPJ)	39/40
Posterior cingulate/precuneus (PCC/PCun)	31/7
Superior temporal sulcus (STS)	21/22
Limbic-paralimbic regions	
Orbitofrontal cortex (OFC)	11/12/47
Ventromedial prefrontal cortex (vmPFC)	10/32
Anterior cingulate/paracingulate cortex (ACC/PrCC)	24/32
Temporal pole (TP)	38
Amygdala	Subcortical
Striatum	Subcortical
Frontal regions	
Dorsomedial prefrontal cortex (DMPFC)	8/9
Dorsolateral prefrontal cortex (DLPFC)	9/46
Inferior lateral frontal cortex (ILFC)	44/45/47

In their review of the neuro-chemical mapping of ToM, Abu-Akel and Shamay-Tsoory (2011) note that the core cognitive ToM network comprises the dorsomedial prefrontal cortex, the dorsal anterior cingulate cortex and the dorsal striatum; see also Figure 4.1 below. More particularly, within the 'mentalizing network', Self-Other representation is supported by dorsal and ventral attention/selection systems at the temporoparietal junction (TPJ) and the anterior cingulate cortex (ACC). Similarly, Carrington and Bailey (2009), in their review of social cognition neuroimaging studies, reported that the MPFC and OFC regions were recruited in over 90% of studies, followed by the TPJ in 58% of studies, ACC in 55%, and STS in 50% of studies.

It should be noted, however, that Abu-Akel and Shamay-Tsoory's neural mapping of ToM, and indeed Carrington and Bailey's (2009), is largely based on data from healthy young adults, and thus there is uncertainty as to whether the regions and networks identified in these reviews is precisely applicable to older adults. This point is important, given that there is a well-documented age-related shift in functional connectivity across pre-frontal, temporal and sub-cortical regions (for a review, see Eyler et al., 2011), all of which are associated with social cognition.

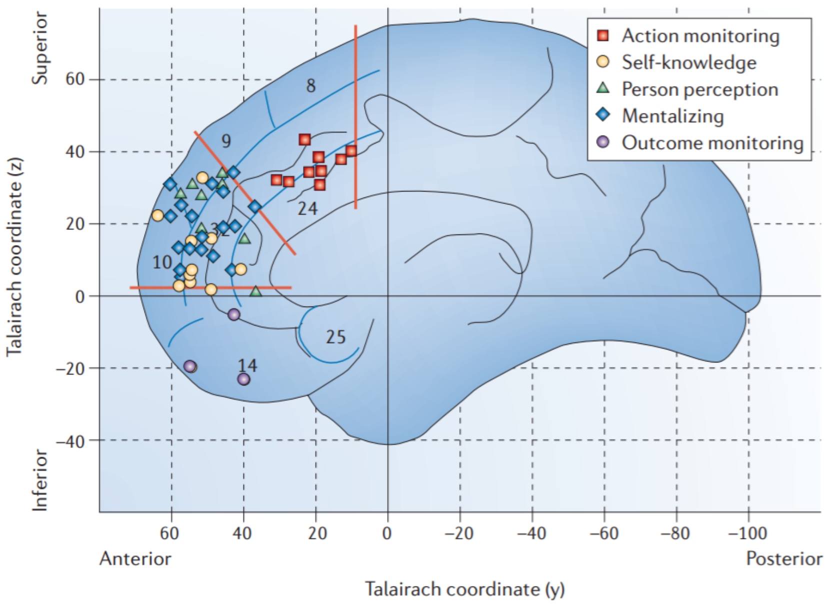


Figure 4.1. Mapping of different aspects of social cognition within the medial frontal cortex. Reproduced with permission from Amodio & Frith (2006).

In differentiating the neural mechanisms involved in modelling the mental states of others versus representing one's own true belief, Vogeley et al. (2001) found that while self-perspective increased activity in the right TPJ and ACC, representing others' mental states increased activity in the ACC and left temporopolar cortex. Building on this, Saxe and Kanwisher (2003) performed fMRI during a ToM task that comprised two conditions: False Belief and False Photo. In each condition, participants read a vignette and were then required to answer a fill-in-the-blank question. In the Belief condition, the vignette and follow-up question pertained to a character's mental or knowledge state, whereas in the Photo condition participants responded to stories on non-human objects or stories that described people in physical detail (as opposed to a description of their mental states). Critically, both conditions require participants to reflect on a situation comprising outdated (false) information, but only the FB condition requires representation of a mental-state. Saxe and Kanwisher found that activation of the TPJ – while present in representing a FB – did not extend to false 'Photo' representations in non-social or control scenarios. Further, greater bilateral blood-oxygen-level dependent (BOLD) activity was found in the TPJ during stories that described characters' mental states compared to stories that described characters in physical detail. This suggests that the TPJ may respond preferentially to mental-state content and not 'falseness' more broadly.

Since their original publication, Saxe and Kanwisher's (2003) ToM experiment has been employed numerous times in the social cognition literature and has been shown to reliably recruit distinct, circumscribed neural regions with robust effect sizes. As a result, this task has been widely used as a *functional localizer* to locate and identify discrete nodes of a core mentalizing network. Broadly, functional localizers are used to constrain analyses of fMRI data to pre-established *functionally specialised* brain regions (e.g., the TPJ for mentalizing or fusiform face area in the case of face perception; see Friston et al., 2006 for detailed overview of functional localizers). Within these localizer-informed, specialised regions, additional and separate analyses of individual task conditions can be scrutinised to examine potential differentiation in recruitment of these areas. Where there is differentiation in these pre-defined regions based on a specific task demand, this can generally be interpreted as the task theoretically conforming (at least on a neural level) to the established patterns and networks of activation within the given cognitive phenomenon. On this basis, Saxe and Kanwisher's (2003) ToM protocol has been used as a functional localizer in a large number of subsequent ToM studies (citation count of original publication, at time of writing > 2000). For recent examples, see Boccadoro et al. (2019), Hughes et al. (2019), and Prochazkova et al. (2018).

As previously noted, the evidence suggests that inhibition of the self-perspective (in order represent the other's FB) might be dissociable from wider EFs and non-social representations. For example, Hartwright et al. (2015) manipulated the salience (high vs. low) of the self-perspective in both mental and non-mental state representation. Differences in activation of the ventrolateral prefrontal cortex (vlPFC) was only found between high and low salience when representing mental state content; there was no vlPFC differentiation for non-mental representation. The implications of this are twofold: firstly, the vlPFC's activity was only modulated by varying mental- and not non- mental state content and second, the effect of the self-perspective in FB reasoning could be differentiated according to its salience. Thus, Hartwright and colleagues' findings particularly highlight prefrontal regions when managing a reality known/unknown manipulation within a FB scenario.

The ageing brain

Literature on the neurophysiology of ageing generally converges into a single, dispirited verdict: that the human brain deteriorates with ageing, even in the absence of pathology. The extent to which ageing has a deleterious effect can be studied and understood through *structural* and/or *functional* methods, and how these approaches map on to changes in cognitive abilities and behaviour in ageing. A structural account of the senescent brain focuses on age-associated neuroanatomical changes primarily through alteration in grey and white matter structures. For example, Persson et al. (2006) conducted a longitudinal analysis on 40 healthy older adults and found significant reduction in hippocampal volume in addition to compromised white matter integrity in the corpus callosum. These age-related changes were also associated with general cognitive decline, in particular, in episodic memory. More globally, numerous MRI studies have shown atrophy (shrinkage) of the cortical ribbon and reduction in cortical thickness in advanced age (Gomori & Melamed, 1985; Madan & Kensinger, 2016; Resnick et al., 2000).

With particular reference to EF, Leong et al. (2017) performed an eight-year longitudinal analysis of structural MRI and neuropsychological measures in 111 healthy older adults with mean age of 67 at baseline. Age-related modification of structural volume was found, with an average of 0.56% total cerebral atrophy and a 3.56% ventricular expansion per year. Importantly, hippocampal atrophy and ventricular expansion was associated with declining EF, as measured through Trail Making, digit span, verbal fluency, and verbal memory tests.

A particular method of indexing structural changes in the ageing brain is through voxel-based morphometry (VBM; see Ashburner & Friston, 2000), which measures differences in neural tissue concentrations based on structural MRI images on a voxel-by-voxel basis. In their seminal study, Good et al. (2001) examined age-related effects on grey and white matter and cerebrospinal fluid through VBM in 465 healthy adults. They reported that global grey – but not white – matter volume declined with age, with the insula, superior parietal gyri, central sulci, and cingulate sulci being most severely affected. There was, however, also a sex difference in the degree of tissue loss in these areas, with males showing significantly greater reductions compared to females.

Functional approaches, on the other hand, look at the temporal synchrony of brain activity within particular defined regions, usually through fMRI. In ageing, compared to healthy younger adults, an increase or reduction of brain activity in older adults is typically interpreted in two ways: *compensation* or *dedifferentiation*, respectively (see Grady, 2012 for a general review of compensation and dedifferentiation in cognitive ageing). The recruitment of increased regional activity (or 'hyperactivity') in older adults, compared with younger adults on a given task, is thought to compensate for functional deficits elsewhere in the brain. More precisely, older adults tend to exhibit greater bilateral activity in frontal regions (i.e., in the PFC) during cognitive tasks in which younger adults show unilateral frontal activation (Reuter-Lorenz & Cappell, 2008). Greater PFC recruitment is thought to compensate for under- or reduced activity in visual processing areas in older adults (also known as the 'posterior-anterior shift in ageing' or PASA; see Davis et al., 2008). In contrast, a general failure to recruit or a reduction of activity in older adults may be indicative of poorer functional specificity in relevant task-related regions (Carp et al., 2011; Park et al., 2004). Diminished laterality and more diffuse patterns of activity have been shown in ageing, with less selective responses to a range of tasks, especially those tapping executive function (for a recent review, see Koen & Rugg, 2019). The reduction in specificity is supportive of the dedifferentiation hypothesis of ageing. A more comprehensive account of neurocognitive ageing is provided in the General Introduction of this thesis (see Chapter 1).

Neural Underpinnings of ToM in Ageing

Though the literature is relatively sparse, evidence for an age-related difference in ToM on a neural level can be seen in structural and functional imaging investigations. Castelli et al. (2010) submitted young and older adults to fMRI while performing the Reading the Mind in the Eyes Task (RMET); a test requiring participants to infer mental states by viewing only the eye region of photographed faces (Baron-Cohen, Wheelwright, Hill, Raste & Plumb, 2001; see Figure 4.2 below for illustration of RMET task).

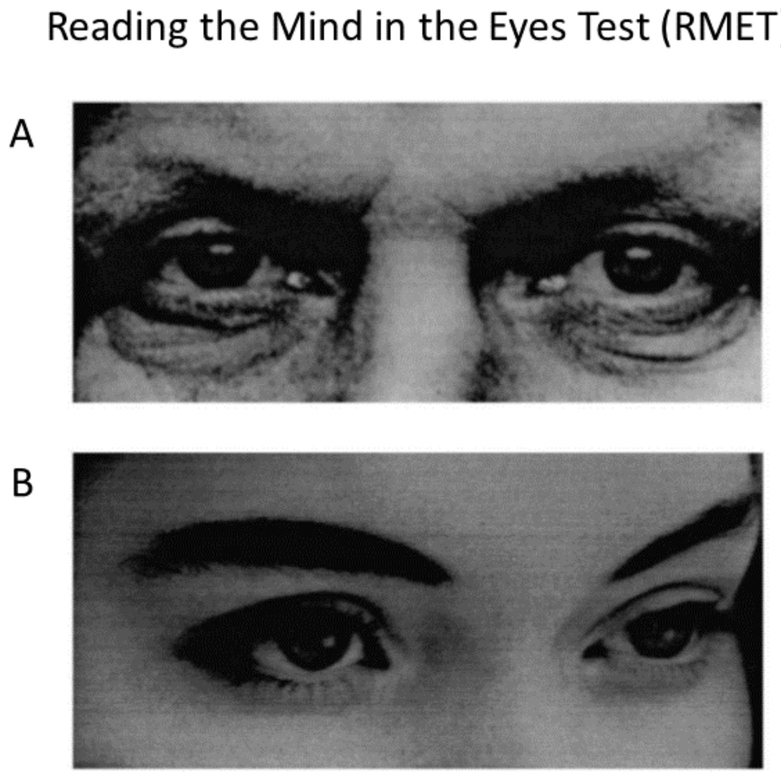


Figure 4.2. Illustration of Reading the Mind in the Eyes Test. From Baron-Cohen et al. (2001). A = illustration of male stimulus. Participants chose from the following options: serious (correct response), ashamed, alarmed, and bewildered. B = illustration of female stimulus. Participants chose from reflective (correct response), aghast, irritated, and impatient.

Castelli and colleagues found no differences in behavioural performance but reported greater activation of frontal regions, especially inferior frontal gyrus (IFG), in older adults. This was contrasted with greater ACC and right frontal gyrus activation in young adults. The RMET was also administered more recently by Cabinio et al. (2015), who measured age-related changes in structural connectivity and its potential association with socio-cognitive abilities in 36 healthy participants. Cabinio and colleagues found that the brain regions susceptible to age-related reduction in white matter connectivity – IFG and superior temporal gyrus – were also the regions most closely associated with mindreading competency as measured with the RMET.

The two previous investigations (Cabinio et al., 2015; Castelli 2010), however, may be somewhat limited in advancing our understanding of age-related changes in *cognitive* ToM. At the time of writing, fewer than four publications have adopted the RMET paradigm when investigating the neural bases of age-related ToM differences. Though there is a more substantial evidence base of the RMET with respect to behavioural differences between young and old (see Kynast et al., 2020 for a recent example), the veracity of the RMET as a 'pure' measure of cognitive ToM has been recently called into question. Oakley et al. (2016) investigated performance on the RMET and an alternative test of ToM – the MASC (Dziobek et al., 2006) – with ASD and alexithymia-matched controls. Alexithymia can be broadly described as the subclinical inability to recognise and articulate one's own and others' emotions – i.e., a general impairment in emotional awareness (for a neuro-coordinate-based meta-analysis of alexithymia, see van der Velde et al., 2013). Oakley and colleagues found that while ASD and alexithymia-matched control participants showed no difference on the RMET, ASD participants showed weakened performance on the MASC. In addition, alexithymia, but not ASD, predicted RMET performance but did not affect MASC performance. Taken together, Oakley et al.'s findings suggest that the RMET may be a better measure of emotion recognition/perception rather than cognitive ToM (see also Quesque & Rossetti (2020) for a more comprehensive discussion and task analysis of cognitive and affective ToM tasks).

Elsewhere, Charlton et al. (2009) studied the association between age-related ToM decline and whole brain volume, white matter hyper-intensities, and white matter integrity, through structural and diffusion tensor imaging. From a sample of 106 older adults aged 50-90, Charlton and colleagues reported that socio-cognitive abilities generally declined with age and this age-related decay was significantly associated with white matter integrity (but not whole-brain volume or white matter hyper-intensities). Finally, Moran, Jolly, and Mitchell (2012) fMRI scanned 31 young and 17 older adults during a variant of Saxe and Kanwisher's (2003) ToM localizer paradigm, in addition to other socio-cognitive tasks. Older adults were marginally more error-prone and slower than younger adults when reasoning about a FB; however, the differences were not statistically significant. This pattern of performance was accompanied by an age-related decline in BOLD signal in one primary region: the DMPFC. More precisely, whereas young adults exhibited significant recruitment of the DMPFC in social (vs. non-social) conditions, older adults showed no differentiation in this region between social and non-social task types, and a lower DMPFC response overall.

More recently, Hughes et al. (2019) examined baseline intrinsic functional connectivity within the core mentalizing network (as described previously) in young and older adults during Saxe and Kanwiser's (2003) ToM localizer task. Overall, Hughes and colleagues reported poorer behavioural performance in the older group, but importantly, weaker connectivity between right TPJ and right temporal pole explained the poorer OA behavioural performance. In summary, there seems to be differentiation in neural activation and connectivity between young and old in ToM in circumscribed, discrete regions. OA show reduced activation of the DMPFC and no differentiation in this region between false vs. true belief conditions, and exhibit weakened connectivity in the TPJ that has been directly associated with poorer behavioural ToM performance. Accordingly, in the current investigation, there are two main candidate regions of focus: the DMPFC and TPJ.

Frontal lobe dysfunction and its relation to ToM in ageing

Volumetric and functional studies of the ageing brain have highlighted frontal regions as being the most susceptible to deterioration from healthy, non-pathological ageing. For example, MR volumetry studies have shown quite clearly a disproportionate reduction in relative cortical volume in frontal areas compared to other regions (Raz et al., 2010; see review by Hedden & Gabrieli, 2004). Further, there is evidence to suggest that there may be a disproportionately greater reduction in regional metabolism in frontal areas: 38% reduction in whole-brain metabolism compared with 42% reduction in the frontal lobe (Tumeh et al., 2007). The accelerated relative degeneration of frontal areas in ageing may in part explain the neural underpinnings of age-related ToM decline, since core cortical nodes implicated in belief reasoning are situated within the frontal lobe (e.g., vmPFC, dlPFC, etc., see Abu-Akel & Shamay-Tsoory, 2011).

However, while we are beginning to understand age-related differences in mental-state reasoning on a neural level, the functional imaging literature to date has seldom factored in age-related changes in neurovascular integrity. Lack of correction for neurovascular variation in young and old may compromise our ability to draw meaningful conclusions about the brain regions involved in the age-related decay of mental-state representation.

The BOLD fMRI signal and neurovascular ageing

The BOLD signal measured by fMRI is a consequence of both neuronal and vascular responses, the latter of which is affected by physiological ageing. The signal itself is an index of the changing concentrations of deoxyhaemoglobin – haemoglobin devoid of oxygen – following neural activity. This activity, which is made up of local synaptic activity and neuronal firing, consumes energy which is replenished or served by increased cerebral blood flow and volume (CBF and CBV, respectively). Broadly, the resultant effect of increased neuronal firing is a transient upsurge in oxygenated haemoglobin which in turn decreases the concentration of deoxyhaemoglobin. As deoxyhaemoglobin and oxyhaemoglobin have differing magnetic properties, fluctuations in their concentration drive BOLD signal variability. Finally, fluxes in the BOLD signal form the basis of the distinctive temporal outline of fMRI: the canonical haemodynamic response function (HRF; for a detailed account of the HRF, see Friston et al., 1998).

More recently, we are beginning to understand that the relationship between neural-vascular coupling and BOLD signal changes with age (Abdelkarim et al., 2019). For example, cerebrovascular reactivity (CVR), a measure of how well intracranial blood vessels control CBF, should be considered an important factor in fMRI analysis, particularly in ageing studies (Tsvetanov et al., 2015). CVR determines the rate of constriction or dilation of local blood vessels in the brain which in turn results in reduced or increased regional CBF, respectively. Fluctuations in regional CBF, for reasons outlined in the previous paragraph, ultimately drive variations in the fMRI BOLD signal.

The integrity of the neurovascular architecture that ‘props up’ the CVR, and indeed CVR itself, has been found to vary as a function of healthy, non-pathological ageing. For example, Miller et al. (2019) scanned young and older adults using 4D flow MRI (which computes intracranial angiographic and quantitative blood flow) to investigate age-related differences in CVR. Overall, Miller and colleagues reported lower global CVR in older adults. Further evidence of CVR decline with age can be seen in Flück et al. (2014), where the authors used a transcranial Doppler ultrasound technique; additionally, similar age-related effects were also reported through a phase contrast imaging method in Geurts et al. (2018). Together, these lines of evidence suggest that ageing compromises the vascular response in meeting metabolic – or neuronal – demands. For these reasons, an interpretation of a straightforward BOLD signal discrepancy between young and older adults – without considering neurovascular integrity – may erroneously inflate any potential brain-based effects of ageing. For a detailed explanation of the dissociation of vascular and neuronal components in BOLD-fMRI, see Tsvetanov, Henson, and Rowe (2021).

Consequently, one must exercise reasonable caution when interpreting the outcomes of the reviewed ageing-ToM literature earlier in this chapter. The imaging ToM work cited thus far has not taken steps to practically correct for or factor in young-old neurovascular changes. To the authors’ knowledge, no fMRI study in ageing-ToM to date has sought to correct for age differences in neurovascular integrity. One such method of adjusting the conventional BOLD response in accounting for age-related differences in vascular integrity is by scaling the fMRI signal with the Resting-State Fluctuation Amplitude (RSFA; see Tsvetanov, 2015). The RSFA method uses fMRI data acquired from participants during a ‘resting-state’. Resting-state imaging is the capture of fMRI during a period of rest, where the participant is not performing any particular tasks. The

BOLD signal that arises from resting-state scans is interpreted as the brain's default – or baseline – state, as it indexes intrinsic and spontaneous fluctuation of brain activity in the absence of an external stimulus (see Biswal et al., 1995). Practically, during resting-state sequences, participants are therefore typically instructed to 'clear' their minds and either focus on a blank screen with a central fixation or keep their eyes closed.

Broadly, RSFA is an index of vascular reactivity and variability that is captured during resting-state fMRI. As such, though not interchangeable, RSFA can be loosely described as resting-state fMRI variability. Since this index is captured in a task-free state, RSFA is therefore more sensitive to vascular rather than neuronal components, and thus is an ideal measure via which task-induced BOLD signal can be scaled to account for neurovascular variability between young and old.

Current Study

In the present study, we built on prior behavioural work (Chapter 3) by manipulating two psychologically relevant parameters within a FB task. We systematically manipulated knowledge of reality (KOR; reality known vs reality unknown) and the congruence of self-other perspectives (SOC; congruent vs incongruent) within a single object-displacement paradigm. In doing so, we separated competition from managing self-knowledge from representing incompatible self-other mental states. In addition to our own ToM paradigm, we employed an extensively-published ToM functional localizer task (Saxe & Kanwisher, 2003) to inform our analysis of fMRI data.

Research Questions and Hypotheses

In the current investigation, as mentioned, our FB task comprises two factors: KOR and SOC. The KOR factor manipulates the presence/absence of knowing the true state of affairs. This factor is essentially non-social. KOR (conceptually identical to the 'KoR' factor in Chapter 3) recreates a common feature of the classic FB task where the participant is informed about the status of a key object that is of interest to the target agent. The SOC manipulation (identical to 'SOC' in Chapter 3), conversely, is social in nature as it systematically varies compatibility in self-other beliefs, where a third-party agent's knowledge state can either be congruent or incongruent with one's own belief. As mentioned previously, classical and contemporary investigations of ageing ToM have produced mixed results due to large methodological heterogeneity in task design, where some paradigms are more or less 'social' than others. Likewise, there is also significant heterogeneity in non-social or executive demands in previous ageing ToM experiments. To examine the potential sources of interference in reasoning or mentalizing about the beliefs of others in older age, we have incorporated these two factors within a single ToM paradigm to clarify the contribution of each in explaining the oft reported age-related decline in ToM. By manipulating these two factors, we could emulate some core task demands seen in FB reasoning and, in doing so, systematically modulate relevant non-social and social brain regions. This would allow us to look at specific patterns of difference in younger versus older adults, in both relevant non-social and social components of FB reasoning.

As such, we expected the previously explained social brain networks to be involved when mentalizing about another agent (SOC), and expected that regions involved in wider domain-general EF processes to be involved in KOR. More precisely, building on from Mars et al.'s (2012) parcellation of the TPJ, where the anterior region was shown to be involved in attentional control mechanisms (see Krall et al., 2015), and the posterior TPJ involved in social cognitive function, we expected that KOR will be associated with greater activation in the anterior TPJ and SOC linked to greater recruitment of the posterior TPJ. In line with the ToM literature, we also expected greater activation of the DMPFC in our SOC (or ToM) condition.

We expected that there would be an age effect in these regions whereby there is a differentiation of patterns of activity between young and old in the TPJ and DMPFC, in line with our theorised social/non-social manipulations. To control for variability in vascular integrity between young and old, we scaled our effects of interest with a measure of neurovascular coupling. We utilised both GLM-based and non-parametric approaches, and performed an unbiased whole-brain analysis which was then followed up with a ToM functional localizer-based region of interest (ROI) assessment.

Adjustment to participant recruitment and data analysis due to COVID-19 disruption

Due to the COVID-19 pandemic and its resultant restrictions, data collection for this study was prematurely terminated. Originally, the anticipated sample for the study was 90 participants – 45 young (YA) and 45 older adults (OA). This was based on a power analysis of 90% statistical power, $\alpha = .05$, and an effect size based on existing open ToM fMRI data from Moran et al. (2012). Sample size estimates were computed using a specialised, web-based fMRI power calculator: NeuroPower (www.neuropowertools.org). However, due to the aforementioned disruption, a total of only 72 participants were recruited – 50 YA and 22 OA. Of the 22 OA, only 19 had complete and usable data (both behavioural and fMRI). Thus, a matched sub-set from the YA data was used in the final analysis, to allow for reasonable between-groups analyses. Matching of the two age groups was performed through two variables: sex and BAPQ scores (explained later in Methods). As such, data from only 38 participants (19 YA, 19 OA) are reported in this chapter. Upon resumption of testing and scanning participants, COVID-19 permitting, the full dataset will be re-analysed and written up separately for publication. It is important to note therefore our anticipated analysis plan was greatly hindered due to low statistical power, and that the findings reported in this chapter should be interpreted with such a context in mind.

Methods

Participants

The study comprised self-declared healthy adult volunteers. Prospective volunteers were asked not to participate if they had any contraindications to MRI, or if they had been diagnosed with any psychiatric, neurological, developmental, or vascular disorder. Thirty-eight (28 female) neurologically healthy participants[3] took part in this study: 19 YA (*M* age = 21.11 years, *SD* = 3.59) and 19 OA (*M* = 69.32 years, *SD* = 4.37). All participants were right-handed (confirmed using the Edinburgh Handedness Inventory Short Form [EHI]; Veale, 2013) and fluent speakers of English, with normal or corrected-to-normal vision. OA were recruited from an institutional participant database and from local community groups. YA were recruited via a student participant pool. Participants were compensated with either course credits (YA) or a payment of cash and gift vouchers totalling 20.00 GBP (OA). The study adhered to the Declaration of Helsinki and was approved by the institutional ethics board (Appendix D). All participants provided written informed consent.

Screening

Suspected Autism – AQ-10

Socio-communicative impairment is characteristic of autistic spectrum disorders, with marked deficiencies in social reasoning. As such, all participants were administered the Autism Quotient-10 (Allison, Auyeung and Baron-Cohen, 2012) to screen for suspected autism. As per Allison e al. (2012), scores of > 6 warrant a referral for further

[3] As previously discussed, the sample size and subsequent analyses of this study were amended following disruptions due to COVID-19.

investigation; thus, participants who scored > 6 were excluded from the analysis. No participants were excluded on this basis.

Dementia – M-ACE

All OAs were screened for dementia using the mini Addenbrooke's Cognitive Examination (M-ACE; Hsieh et al., 2015). The M-ACE has a lower cut-off at 21 (/30) which may be indicative of dementia. Consequently, OAs scoring < 21 were excluded from the data analysis. One older participant was excluded for scoring below 21 on the M-ACE.

Contraindication to MR imaging

All participants were screened for suitability of MRI scanning through an institutional MRI safety questionnaire (Appendix E), performed by the on-duty research radiographer. Participants who showed contraindications (to health and safety and/or data quality) were excluded and were not permitted to continue with the study. No participants were excluded on this basis.

Incidental findings

The current study adhered to an institutional protocol for the detection and reporting of incidental findings. Where incidental findings were confirmed by a physician as clinically relevant, such participants were excluded from the final analyses. All images were initially screened for suspected incidental findings by the on-duty research radiographer.

Design

False belief task

The ToM paradigm was an iteration of our previous behavioural work (Chapter Three). The task was a two-factor (2 x 2) design wherein each factor manipulated a potential source of conflict or interference; see Table 4.2 and Figure 4.3.

Table 4.2.

Summary of Experimental Factors and Levels

		Factor levels	
Condition	Factor descriptor	0	1
Knowledge of Reality (KOR)	Manipulates whether the participant is given explicit knowledge about the true state of affairs	reality unknown (KOR_0)	reality known (KOR_1)
Self-Other Congruence (SOC)	Manipulates the congruence of the participant's and the target agent's beliefs about what is the state of affairs	self-other congruent (SOC_0)	self-other incongruent (SOC_1)

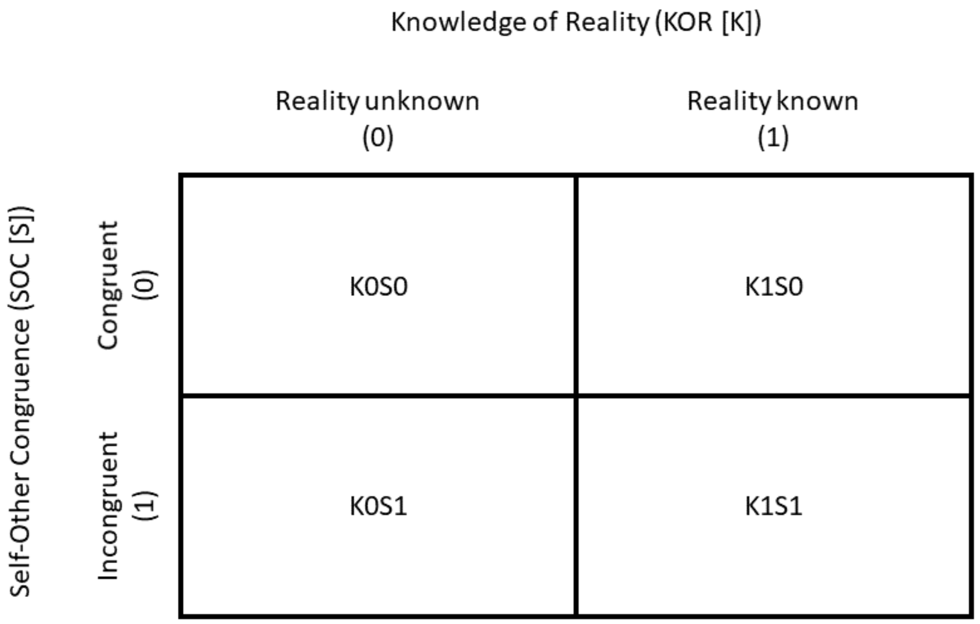


Figure 4.3. *Factor structure. Each quadrant represents a single experimental condition.*

The first factor, termed ‘Knowledge of Reality’ (KOR), systematically varied the participant’s explicit knowledge of the true state of affairs, informed by previous work suggesting that one’s own true, self-knowledge may interfere with belief reasoning when representing a false belief; for example, refer to Birch & Bloom’s (2004; 2007) ‘curse of knowledge’ hypothesis. Thus, the KOR factor comprised two levels: reality unknown (KoR0) and reality known (KoR1). The second factor, ‘Self-Other Congruence’ (SOC), manipulated the congruence in belief states between participant and agent. We have previously shown that a mismatch in self-other belief states is a source of interference in ToM reasoning, and this interference seemed to disproportionately affect older adults (Chapter Three). The SOC factor therefore also comprised two levels: self-other congruent (SOC0; compatible self-other beliefs) and self-other incongruent (SOC1; incompatible self-other beliefs). Overall, the task consisted of four experimental conditions (see Figure 4.3). Each experimental trial depicted a computerised ‘game’ wherein the participant was presented with two cups, one of which contained a target (a red ball) and the other was empty. Simultaneously, the participant saw an on-screen character (herein known as the ‘agent’) who also played game. The agent indicated the cup they believed to contain the ball by resting their finger above the relevant cup. Critically, the participant received a clue – which the agent was not privy to – by way of a single cup being revealed. The revealed cup was either empty or contained the target ball. Finally, a probe was presented which required the participant to either resolve the trial via the representation of a belief state or from the participants own true belief. The probe alternated between THINK, REALLY, and EMPTY. The THINK probe required the participant to respond with where the agent believed the ball was (ToM trial); the REALLY probe required the participant to respond with which cup actually contained the ball (non-ToM trial); and finally the EMPTY probe required the participant to respond with which cup was empty (non-ToM trial). The latter two variants were non-ToM trials in that they could be resolved from the participant's own true belief, without the need of representing – or mentalizing – the agent's mental state. This manipulation between ToM and non-ToM trials permitted differentiation between true and false beliefs, and further confirmed whether participants were attending to the clue and appropriately engaging with the task.

Each trial comprised seven images with the first image always depicting the agent and two occluded cups. The final image of the sequence always presented the response probe, with the three aforementioned cues (i.e., THINK, REALLY & EMPTY). The intermediate images comprised a clue – where one cup was revealed – and a depiction of the agent's belief of the location of the target; to increase engagement and understanding of the task there were also textual prompts in-between the critical events (see Figure 4.4).

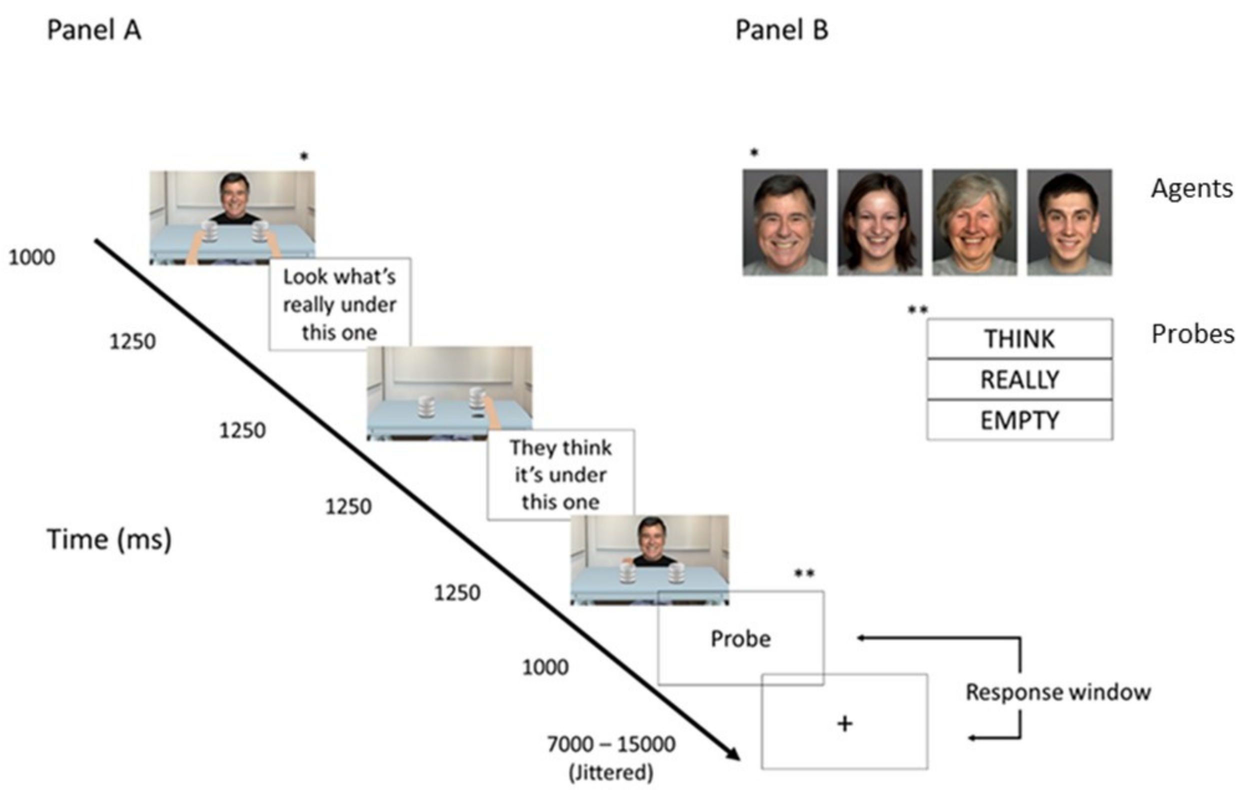


Figure 4.4. Task outline and timings. Panel A: Time course and events of a typical trial. Panel B: Variants of experimental events. The order of all events, save the opening fixation cross and the end probe, varied from trial to trial, counterbalanced both within and across experimental blocks.* Agent varied across trials – counterbalanced across experiment to include young/old male/female; see Panel B for images of all agents. ** Probe varied across trials – see Panel B for three possible variants. The THINK probe required the participant to respond with where the on-screen agent thought the ball was, the REALLY probe required the participant to respond with the ball actually was, using their own self-knowledge, and finally the EMPTY probe asked participants to indicate which of the two cups was empty (or did not contain the target ball).

The agent (Figure 4.4, Panel B) in the task varied between male and female, and young and old; and there were therefore four possible agents: young male, young female, older male, and older female. The presentation of agents was counterbalanced across conditions, which meant that there were equal presentations of young/old male/female agents in each of the four experimental conditions. An important consideration here is potential unintended interference from the avatars themselves (i.e., due to age-incongruent agents). For example, Ferguson, Brunsdon, and Bradford (2018) reported differential altercentric bias from age-in/congruent avatars in a visual perspective-taking context; where, overall, there seems to be increased processing or cost of mentalizing for own-age others. In light of this, we attempted to reduce any potential interference from age-in/congruent avatars in the current design by exposing all participants (young and old) equally to ostensibly young and older avatars (or agents).

The original task on which this current iteration is based, had three factors, which included an Other-Other Congruence manipulation based on dis-/agreement in beliefs of two third-party agents. This factor highlighted a particular confound with respect to the cuing and subsequent processing of multiple locations that is a typical feature in many false belief paradigms. The cuing of these multiple and often competing locations tapped executive resources – particularly spatial working memory – that were typically decayed in older adults; thus confounding any actual socio-cognitive differences. We showed, however, from a precise experimental manipulation where the number of cued locations was kept the same and the only difference between the conditions was resolving a false (vs. true) belief, that reasoning about a false belief was effortful above and beyond processing multiple cued locations. For this reason, we did not include the Other-Other Congruence manipulation in the current design.

Theory of mind localizer

A variation of Saxe and Kanwisher's (2003) ToM functional localizer was used to identify discrete nodes of the mentalizing network to inform analyses of our novel false belief paradigm. For a detailed task analysis and breakdown of the ToM localizer, see Dodell-Feder et al. (2011). The stimuli consisted of 32 stories equally split into two conditions: 1) Belief, and 2) Photo. Both conditions required participants to represent and respond to false or outdated content, the essential difference between the two conditions was the content being represented: beliefs or photographs/maps. The Belief condition required participants to mentalize about a protagonist's belief state whereas the Photo condition could be resolved using physical or mechanical inferences. The stories were then proceeded by a true or false question. An example of a Belief trial is as follows: *Anne made lasagne in the blue dish. After Anne left, Ian came home and ate the lasagne. Then he filled the blue dish with spaghetti and replaced it in the fridge*; with the subsequent follow-up: *Anne thinks the blue dish contains spaghetti* (participants then asked to respond with 'True' or 'False'). An example of a Photo trial is as follows: *A photograph was taken of an apple hanging on a tree branch. The film took half an hour to develop. In the meantime, a strong wind blew the apple to the ground*; followed-up with: *The developed photograph shows the apple on the ground*: True or False.

Participant response-time (RT) and accuracy data were recorded. Each condition was represented equally (16 Belief, 16 Photo trials) across two experimental blocks, with conditions counterbalanced across runs.

Questionnaire Measures

Broad autism phenotypes

The Broad Autism Phenotype Questionnaire (BAPQ; Hurley et al., 2007) was administered to examine participants' phenotypic expression of autism. The BAPQ comprises three subscales: social aloofness, rigid personality, and pragmatic language – and is designed to survey sub-clinical autism traits in non-autistic adults (Hurley et al., 2007). The psychometric properties of the BAPQ have been reported to be moderately-high to high, with the following internal consistency scores (Cronbach α): Social aloofness = .92, Rigid personality = .86, and Pragmatic language = .80 (Sasson et al., 2013).

Reading competency

Participants' word reading ability was measured using reading component of the Wide Range Achievement Test, Third Edition (WRAT-3; Snelbaker, Wilkinson, Robertson & Glutting, 2001).

Experimental Procedure

Questionnaire measures and practice session prior to fMRI

After informed consent (and prior to the MRI scanning session), participants completed the questionnaire and screening measures in the following fixed order: AQ-10, EHI, BAPQ, WRAT-3, and M-ACE (OA only). After completion of these self-report measures, participants underwent a ~30-minute training session which comprised a self-paced Microsoft PowerPoint presentation that outlined the nature of the novel ToM task and the localizer task. After the PowerPoint, participants had the opportunity to practice the tasks via two pre-programmed PsychoPy (Peirce et al., 2019) experiments – one for each task. While the content of the trials differed in the practice sessions to that of the actual study experiment, the technical parameters of the trials – e.g., timings, frequency of conditions presented, stimuli dimensions, etc. – remained the same. Participants were allowed to repeat the PowerPoint presentation and/or the practice experiments as many times as they wished – though, in practice, no participant required more than two practice runs. Completion of the questionnaires and practice session typically took around one hour and then participants were moved to the MRI facility to begin scanning.

f/MRI Procedure

Participant brain data was recorded in the following order: T1 anatomical, fMRI false belief, fMRI ToM localizer, and an eyes-open resting state; see Figure 4.5 for typical order of f/MRI sequences. The trial structure of each ToM task is described separately below.

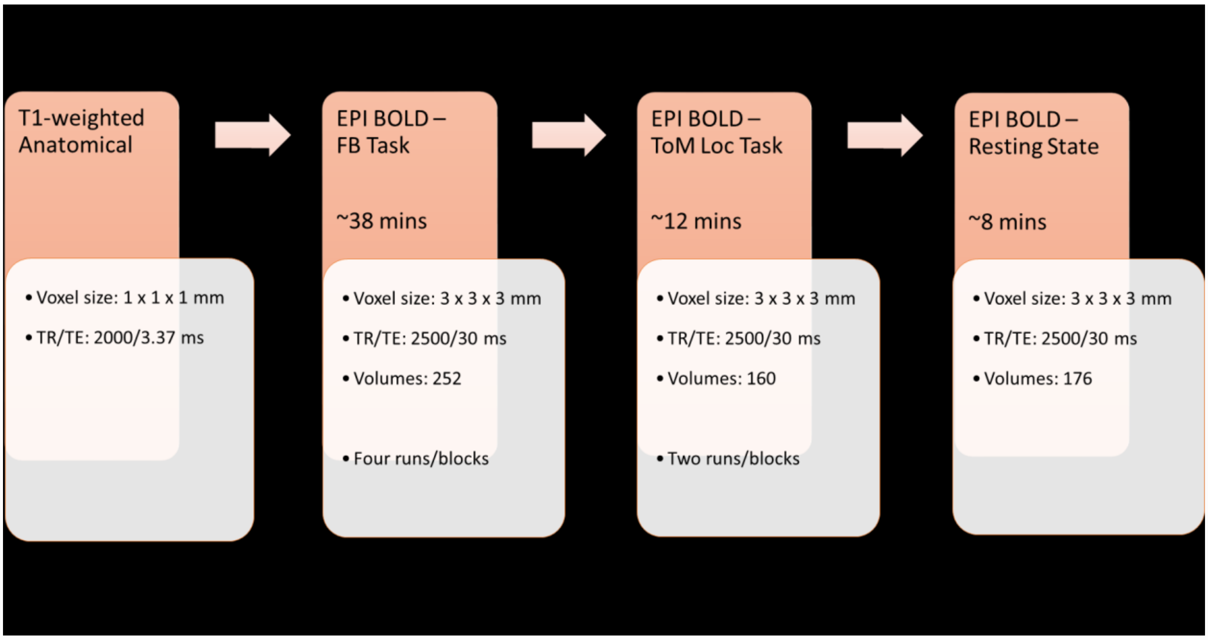


Figure 4.5. *Typical order of sequences and their respective scanning parameters.* EPI = Echo Planar Imaging. BOLD = Blood-oxygenated-level-dependent. FB = False Belief. ToM Loc = Theory of mind localizer. TR = Repetition time. TE = Echo time.

fMRI tasks. All fMRI tasks were presented using E-Prime (v 2.0.10, Psychology Software Tools). The stimuli were projected from a Lenovo Intel i7 Windows 10 Pro ThinkCentre stimulus computer outside of the MR scanner room. The stimuli were projected from a rear-projection screen and the participant was able to view these stimuli via a mirror mounted on the head coil. Responses to stimuli, if required, were captured using a Cedrus Lumina MR-compatible left/right button response box placed in the participant's right hand (all participants were right handed).

False belief task. Each experimental trial comprised seven events, beginning with a 1000 ms presentation of an opener (a slide depicting one of the four possible agents sitting behind a desk; see Figure 4.4). The four subsequent events, all presented for 1250 ms, provided textual prompts to the participant, displayed the content of a single cup, and indicated the agent's belief with respect to their chosen cup. Following on from this, a 1000 ms probe (Think, Really, Empty) was presented to the participant. Trials ended with a variable – or jittered – inter-trial-interval (ITI) that ranged 7000 – 15000 ms, increasing in 1000 ms increments (mean ITI across all trials = 11000 ms). ITI iterations were counterbalanced across conditions and runs. Participants responded by pressing the appropriate left/right button as depicted from their perspective (thus; a left response corresponded to the left cup) Participant responses were captured from the onset of the probe slide until the end of the ITI; participant responses did not prematurely terminate the ITI. Thus, on average, a trial lasted 18 s (critical events = 7000 ms, mean ITI = 11000 ms).

In total, there were 128 trials split into four blocks, each block comprising a single fMRI run. Each block therefore contained 32 trials – 16 ToM (THINK) trials and 16 non-ToM (REALLY / EMPTY) trials. Experimental condition, ToM/non-ToM trials, presence/absence of target ball in each cup, task agents, 'correct' left/right response, and duration of jittered ITI were counterbalanced within and across fMRI/experimental blocks. Each block lasted 9 min and 36 s (576 s), and the whole task, not including self-timed breaks in between blocks, lasted approximately 38 min and 24 s (2304 s).

Theory of mind localizer. Trial design and timing of events for Saxe and Kanwisher's (2003) ToM localizer task were taken from Dodell-Feder et al. (2011). Each trial began with a 6 s presentation of a story, immediately followed by a 4 s question wherein participants were required to make a True or False response (left/right on button-box). Finally, a 10 s ITI was presented (where button responses were still allowable). Thus, a single trial lasted 20 s. The task comprised 32 trials in total split into two 16-trial blocks, each block comprising a separate fMRI run. Each block therefore lasted 5 min and 20 s, and the total experiment – excluding self-timed breaks – took around 10 min and 40 s.

Resting state. Participants underwent an eyes-open resting state sequence where they were instructed to concentrate on a central white fixation cross projected onto a black background. Participants were instructed to keep their eyes open, avoid falling asleep and to 'clear their mind'. The resting state sequence lasted 7 min and 30 s.

Timing of overall study. The entire scanning session – an anatomical scan, two task fMRIs, and a resting state scan – took approximately 65 – 70 min. Depending on length of comfort and self-timed breaks, the entire study (pre-scan questionnaire and practice period combined with imaging session) took around 130-150 min.

Image acquisition

Acquisition and pre-processing parameters are reported according to the Organization of Human Brain Mapping's Committee on Best Practice in Data Analysis and Sharing (COBIDAS) guidance (see Nichols et al., 2016).

A 3T Trio (Prisma) Siemens scanner at the Aston Brain Centre, Birmingham, UK was used to obtain structural and functional data. All images were obtained with a 32-channel (Siemens Head 32, Siemens Corp) head coil. High-resolution anatomical scans were obtained via a 1 mm isotropic T1-weighted sequence (TR = 2000 ms, TE = 3.37 ms, flip angle = 8º; 1 mm slice thickness, no gap) while participants were looking at a blank, black screen. Single-shot echo-planar imaging (EPI) procedures were used to capture BOLD signal during both task and resting state scans. The imaging parameters for each task/sequence were as follows. False belief: TR = 2500 ms, TE = 30 ms, flip angle = 83º, 3 mm slice thickness, FOV = 250 mm, and bandwidth = 1130 Hz/Px. ToM localizer: TR = 2500 ms, TE = 30 ms, flip angle = 90º, 3 mm slice thickness, FOV = 240 mm, and bandwidth = 1078 Hz/Px. Resting state: TR = 2500 ms, TE = 30 ms, flip angle = 80°, 3 mm slice thickness, FOV = 240 mm, and bandwidth = 1078 Hz/Px. Functional images were acquired in an interleaved slice order with a voxel size of 3 x 3 x 3 mm. Partial brain coverage was achieved in the functional scans (to increase spatial resolution/hasten acquisition time) by omitting the brainstem and lower half of cerebellum.

Data pre-processing and analysis set-up

Behavioural data

RT and accuracy data (recorded in the scanner via E-Prime) were first merged using the E-Merge application (Psychology Software Tools Inc, PA, USA); merged data were then transferred into an Excel spreadsheet. From Excel, the behavioural data were imported into various statistical packages, including JASP, SPSS, and R.

Neuroimaging data

Processing and cleaning of neuroimaging data was performed via FSL (v. 6.0.1; Oxford Centre for Functional MRI of the Brain, University of Oxford) and MATLAB (v. R2020a; MathWorks, USA). For a brief overview of FSL tools and neuroimaging applications used in the current pre-processing and analysis pipeline, see Table 4.3.

Table 4.3.
Tools and imaging applications used in pre-processing, visualising, and analysing fMRI data

Tool	Purpose
BET/BET2	Extraction of brain from non-brain tissue (FSL)
FLIRT	Linear brain image registration (FSL)
MCFLIRT	Extension to FLIRT; motion correct tool for fMRI time series (FSL)
FEAT	Model-based analysis of fMRI data (FSL)
MELODIC	Independent Components Analysis to decompose fMRI data into 'real' task-related signal or artefactual components (FSL)
SOCK	Software toolbox used to identify artefact components following ICA (MATLAB)
Randomise	Non-parametric, permutation-based inference/thresholding of fMRI statistical maps (FSL)

Source: https://fsl.fmrib.ox.ac.uk/fsl/fslwiki; https://www.nitrc.org/projects/sock/

First, brain extraction was performed using the BET utility in FSL with a common centre-of-gravity used for all participants, and a default fractional intensity threshold of 0.5. A manual visual examination of both the resulting extraction and registration (i.e., high-resolution anatomical and functional images) was performed to assess accuracy of these processes. Slice timing correction was performed with the *regular up* option and data were spatially smoothed using a Gaussian filter (full-width half maximum = 5 mm). A default high-pass filter was applied. The standard approach to motion correction was undertaken in MCFLIRT, which revealed that no participant exceeded 1 mm of mean relative displacement, for either the FB or localizer tasks. In terms of registration, the middle volume of each participant's functional scan was registered to their respective high-resolution T1 anatomical image; the anatomical image was then registered to a common brain space (MNI template, Montreal Neurological Institute). We further examined artefactual signals by performing an independent components analysis (ICA) in MELODIC to identify signal related to residual head motion and physiological noise (e.g., heart rate, respiration, white matter, cerebrospinal fluid). These artefactual components were identified using SOCK (Spatially Organized Component Klassifikator, see Bhaganagarapu, Jackson, and Abbott, 2013) - an automated classifier that separates artefact-related signals from those that are likely to be of neuronal origin. Finally, these artefuactual components were removed (i.e., regressed out) from the functional time series.

Analysis set-up. Analysis was undertaken in FSL FEAT, where explanatory variables (EVs) were created each representing a single task condition; for the FB task, there were four EVs (K0S0, K0S1, K1S0, K1S1), and two EVs for the ToM localizer (Belief, Photo). Each EV was convolved with the standard double-gamma HRF within a GLM framework. For both tasks, modelling of participant brain responses comprised the duration between the onset of the experimental protocol (trial) and the participant button-response. Second and third level modelling comprised the aggregation of session data per task, per participant, and then the data were aggregated across participants using a random effects model.

The modelling of the contrasts in the FB task was guided by our previous behavioural findings (i.e., Chapter Three). See Table 4.4 for details of each contrast.

Table 4.4.
Overview of contrasts in the analysis of task-fMRI data

Contrast	Note
Main Effects	
KOR 0 > KOR 1	Main effect of Knowledge of Reality
SOC 1 > SOC 0	Main effect of Self-Other Congruence
Interaction Terms	
SOC 0: KOR 0 > KOR 1	Effect of Knowledge of Reality in the SOC Congruent (SOC_0) condition
KOR 1: SOC 1 > SOC 0	Effect of Self-Other Congruence in the KOR 'Reality Known' (KOR_1) condition
ToM Localizer	
Belief > Photo (B > P)	Effect of the mentalizing condition (Belief) over non-social condition (Photo)

KoR = Knowledge of Reality (0 = unknown; 1 = known). SOC = Self-Other Congruence (0 = congruent; 1 = incongruent). OOC = Other-Other Congruence (0 = congruent; 1 = incongruent).

The KOR 0 > KOR 1 contrast was modelled as such to reflect our previous finding that a knowledge of reality (i.e., KOR 1) enhanced false belief performance, while the SOC 1 > SOC 0 reflects findings that indicate an incongruence in self-other beliefs (SOC 1) hinders or interferes with false belief reasoning (Chapter Three). Each of these contrasts were then further scrutinised on a between-groups (YA vs. OA) basis. GLM-based random-effect analyses were performed to examine potential group-level effects based on our proposed contrasts. Parametric maps resulting from the GLM analyses were then corrected for multiple comparisons with non-parametric permutation testing. A default of 5000 permutations were run for each non-parametric test.

Region-of-interest analyses. ROI analyses were based on the ToM localizer. ROI masks comprised spheres of 10 mm radius centred on group-level peak voxels highlighted in the Belief > Photo contrast. Only clusters with a peak Z value of > 4 were considered and this resulted in the extraction of six ROIs/clusters. Within these clusters, mean COPE values for each predefined FB contrast was extracted using the Featquery tool in FSL. COPE images/values are calculated by subtracting the parameter estimate for EV 1 (explanatory variable, e.g., 'KOR 0') from EV 2 (e.g., 'KOR 1'). Thus, the larger the difference between the two EVs the larger the magnitude of the effect or contrast within this given test. Where there are many EVs, a contrast is created by modelling each input as 1, -1, or 0, depending on the variables of interest, where inputs coded as '0' remain constant across all levels of the other variables. Extracted COPE values were then analysed through a series of four – one per contrast – mixed ANOVAs in JASP, where Age was entered as a between-subjects factor and regions (as highlighted by ToM localizer) as within-subjects.

Calculation of RSFA index. Calculation of RSFA was based on Tsvetanov et al. (2015). RSFA was calculated by computing the standard deviation of each brain voxel over the course of the resting-state scan, within a standard grey matter mask. This resulted in an individual three-dimensional RSFA map for each participant. These maps were then scaled (or normed) by dividing each voxel value by the sample maximum. These values were then used to scale the resulting contrast images from the GLM analysis, on a participant-by-participant basis. Only data from the FB task was scaled with RSFA; ToM localizer data was not scaled in order to recruit ROIs consistent with the prior literature. To our knowledge, no studies to date have scaled ToM localizer functional data with RSFA when comparing young vs. older adults. However, this is an area of interest which will be pursued when data collection can be completed (which will commence once the COVID situation permits).

Results

Behavioural analysis

Overall, the two age groups were well-matched on sex ratio, highest level of education, and self-reported BAPQ scores (see Table 4.5). As is consistent with the ageing literature, OA showed greater reading ability (WRAT-3) compared to their younger counterparts.

Table 4.5.
Sample characteristics by group, including inferential test statistics

		YA (N = 19)	OA (N = 19)	Test statistic (t / X^2)
Age, years	*M* (*SD*), range	21.11 (3.59) 18 – 29	69.32 (4.37) 62 – 77	t = 37.15***
Sex	Male : Female	5 : 14 (26% : 74%)	5 : 14 (26% : 74%)	X^2 = 0.00
Highest Education	Compulsory/FE	11 (58%)	7 (37%)	X^2 = 1.69
	First/Higher Degree	8 (42%)	12 (63%)	
WRAT-3, total	*M* (*SD*), range	50.26 (5.09) 38 – 55	54.32 (2.81) 48 – 57	t = 3.04**
BAPQ, total	*M* (*SD*), range	85.58 (11.64) 71 – 116	85.16 (17.37) 48 – 122	t = .09

Note: YA = Young Adults. OA = Older Adults. *** p < .001. ** p < .01. * p < .05.
FE (Education) = Further Education. WRAT = Wide Ranging Achievement Test (Reading Component).
BAPQ = Broad Autism Phenotype Questionnaire.

Data screening, processing, and cleaning

Prior to performing confirmatory statistical analysis of the behavioural data, the data were pre-processed. First, to test whether participants were effectively engaging with the task, 'filler' ('Empty' and 'Really') trials were examined to highlight participants who had scored below chance (< 32 based on a binomial probability distribution, $p < .05$). All participants scored above chance and were therefore included in the confirmatory analysis. Subsequently, only ToM ('Think') trials, where a correct response was given, were analysed. For behavioural analyses, incorrect responses on ToM trials, including null responses, were omitted – this resulted in 189 trials being removed (86 and 103 in YA and OA groups, respectively; see Table 4.6 below) prior to analysis. For the functional imaging analysis, incorrect and null responses were modelled as zero and the variability from these trials was not considered.

Table 4.6.
Trials removed per condition as a function of age group, due to incorrect responses or omissions.

	K0S0	K0S1	K1S0	K1S1	Total
YA	27	24	18	17	86
OA	19	36	17	31	103
Total	46	60	35	48	189

YA = Younger adult. OA = Older adult.

False belief (Magic Cups) task RT and error-rate

The structure of the current analyses for the behavioural (and indeed fMRI) data was modelled on the hypotheses and findings from Chapter 3. Primarily, the behavioural hypotheses were examined using a series of factorial analysis of variance (ANOVA) tests on both accuracy (error rate) and RT data. Mean accuracy/RT data were entered into mixed ANOVAs where the experimental conditions served as within-subjects factors and age group was the sole between-subjects factor. To check for similarity in variance in accuracy and RT across both age cohorts, a series of Levene's tests were computed. These showed that the two groups shared comparable levels of variation across all experimental conditions, in accuracy and RT (all Levene's p's > .05).

Accuracy data. First, accuracy data were examined. A three-way (2 x 2 x 2) mixed ANOVA, with KOR (reality unknown vs. reality known) and SOC (congruent vs. incongruent) as within-subjects factors and Age Group (YA vs. OA) as a between-subjects factor, was first performed on the accuracy data. See Figure 4.6 for breakdown of mean error-rates and RTs across the four conditions.

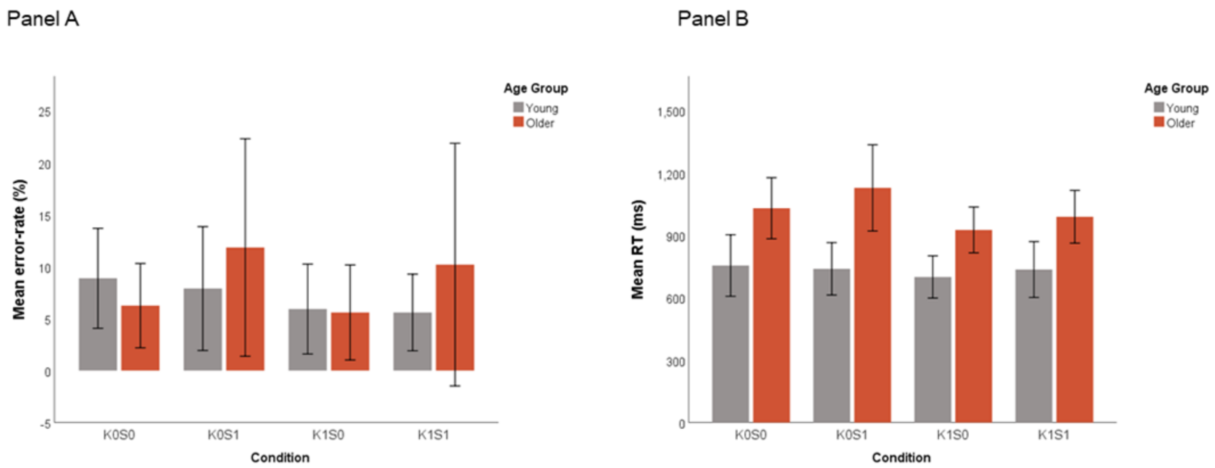


Figure 4.6. *Mean error-rates and RTs across experimental conditions, separated by age group. Error bars represent standard error. K0 = Reality Unknown. K1 = Reality Known. S0 = Congruent self-other beliefs. S1 = Incongruent self-other beliefs. Grey bars represent young adults and orange bars represent older adults.*

The analysis revealed no main effect of Age Group on accuracy. Similarly, there were no effects of KOR and/or SOC. Also, there were no significant interactions between the above-mentioned effects. See Table 4.7 below for full analysis of variance (ANOVA) results.

Table 4.7.
Results of mixed ANOVA on accuracy data

Effect	df	*F*	*p*	Partial ηp^2
Between-Subjects				
Age	1, 36	.114	.737	.003
Within-Subjects				
KOR	1, 36	3.506	.069	.089
SOC	1, 36	1.160	.289	.031
Interaction Effects				
Age x KOR	1, 36	.537	.468	.015
Age x SOC	1, 36	1.949	.171	.051
KOR x SOC	1, 36	.007	.935	.000

Note: KOR = Knowledge of Reality. SOC = Self-Other Congruence.

Response-time data. Next, RT data were entered into a mixed ANOVA as described above. As expected, there was an overall main effect of Age Group where, on average, YA were ~287 ms faster compared to OA, F(1, 36) = 10.373, p = .003, np2 = .224. The analysis also revealed a main effect of KOR, where knowing reality (vs. not knowing/reality unknown) enhanced performance by around 75 ms, F(1, 36) = 11.061, p = .002, np2 = .235. This facilitatory effect of knowing reality – that a knowledge of reality is associated with faster response times – is a replication of a previous finding in Chapter 3. See Table 4.8 below for full analysis of variance results.

Table 4.8.
Results of mixed ANOVA on RT data

Effect	df	*F*	*p*	Partial np^2
Between-Subjects				
Age	1, 36	10.373	.003	.224
Within-Subjects				
KOR	1, 36	11.061	.002	.235
SOC	1, 36	1.996	.166	.053
Interaction Effects				
Age x KOR	1, 36	4.154	.049	.103
Age x SOC	1, 36	1.217	.277	.033
KOR x SOC	1, 36	.038	.847	.001

Note: KOR = Knowledge of Reality. SOC = Self-Other Congruence.

There was a statistically significant interaction between Age Group and KOR, $F(1, 36) = 4.154$, $p = .049$, $np^2 = .103$. To further scrutinise this interaction, follow-up pairwise comparisons were performed and Bonferroni corrections were applied to address increased family-wise error. Figure 4.7 below illustrates the interaction between Age and KOR.

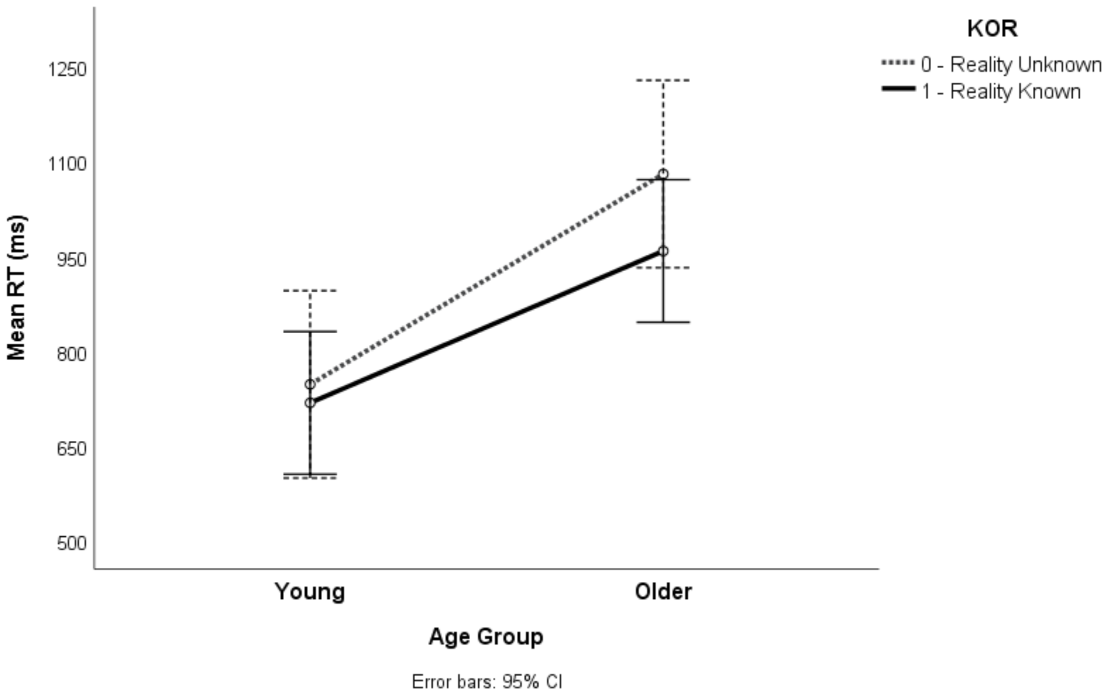


Figure 4.7. *Interaction between Age Group and KOR. Error bars represent 95% confidence intervals.*

In YA, there was no difference between the two levels of KOR ($M_{difference}$ = 29.223), p = .369; see Table 4.9 for full results of pairwise comparisons. However, there was a significant difference between the two levels of KOR in the OA group ($M_{difference}$ = 121.714), p = .001. Accordingly, the main effect of KOR was largely driven by the older cohort, where OA were significantly slower when reality was unknown versus known. Further, when examining the effect of age on each level of KOR, both comparisons returned statistically significant effects (both p's < .01).

Table 4.9.
Pairwise comparisons comparing each level of KOR with Age, and Age with KOR.

	M RT	*M* RT	*M* Difference	*p*	95% CI for Difference
	KOR 0	KOR 1			
Young	747.54	718.31	29.22	.369	-35.86 – 94.31
Older	1080.17	958.45	121.72	.001	56.63 – 186.80
	Young	Older			
KOR 0	747.54	1080.17	-332.63	.003	-542.03 – -123.23
KOR 1	718.31	958.45	-240.14	.004	-399.60 – -80.68

Note: KOR 0 = Knowledge of Reality, Reality Unknown. KOR 1 = Knowledge of Reality, Reality Known.

Finally, correlational analyses between questionnaire data (WRAT and BAPQ sub-scales) and experimental conditions (RT and accuracy) were run. See Tables 4.10 and 4.11 for output of correlation analysis. As the WRAT and BAPQ data did not correlate with any of the experimental manipulations, no further statistical investigations were performed with these two questionnaire measures.

Table 4.10.

Correlation of experimental conditions (RT) and questionnaire measures.

	1	2	3	4	5	6	7	8
1. WRAT	--							
2. BAPQ_Aloof	0.022	--						
3.BAPQ_PragLang	0.243	0.116	--					
4. BAPQ_Rigid	-0.137	.402*	0.085	--				
5. K0S0	-0.067	-0.094	-0.023	-0.249	--			
6. K0S1	-0.034	-0.158	-0.006	-0.264	.734**	--		
7. K1S0	-0.094	-0.154	-0.039	-.344*	.924**	.678**	--	
8. K1S1	-0.060	-0.061	-0.014	-0.196	.888**	.822**	.807**	--

Note: WRAT = Wide Range Achievement Test (only Reading component administered). BAPQ = Broad Autism Phenotype Questionnaire, comprising three subscales; Aloofness, Pragmatic language, and Rigidity. Pearson's correlation. *. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed). 'K' represents Knowledge of Reality factor, and 'S' represents Self-Other Congruence. '0' represents low conflict manipulation and '1' represents high conflict.

Table 4.11.

Correlations of experimental conditions (accuracy) and questionnaire measures.

	1	2	3	4	5	6	7	8
1. WRAT	--							
2. BAPQ_aloof	0.022	--						
3. BAPQ_pragLang	0.243	0.116	--					
4. BAPQ_rigid	-0.137	.402*	0.085	--				
5. K0S0	-.244	-0.020	-0.044	0.059	--			
6. K0S1	-0.199	-0.105	-0.101	0.017	.643**	--		
7. K1S0	-0.206	0.034	-0.007	0.157	.818**	.735**	--	
8. K1S1	-0.006	-0.020	0.057	0.094	.576**	.832**	.675**	--

Note: WRAT = Wide Range Achievement Test (only Reading component administered). BAPQ = Broad Autism Phenotype Questionnaire, comprising three subscales; Aloofness, Pragmatic language, and Rigidity. Pearson's correlation. *. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed). 'K' represents Knowledge of Reality factor, and 'S' represents Self-Other Congruence. '0' represents low conflict manipulation and '1' represents high conflict.

Practice/fatigue effects

Due to the length of the false belief task, additional analyses were performed to examine if there were disproportionate effects of learning or fatigue between the two cohorts. Thus, two (for accuracy and RT, separately) further 4 x 2 mixed ANOVAs with Experimental Block (1, 2, 3, and 4) and Age Group (young vs. older), were performed. The analyses revealed no significant interaction between Experimental Block and Age Group, for either accuracy or RT data, (both p's $> .05$); which suggested that there were no significant effects of learning/fatigue between the two groups.

ToM Localizer

To examine the behavioural ToM localizer data, two separate mixed ANVOAs were performed, on accuracy and RT, respectively. Accordingly, age was entered as a between-groups factor and Condition (Belief vs. Photo) was entered as a within-subjects factor in the analysis. Mean condition error-rate and RT, as a function of Age Group, can be seen in Figure 4.8.

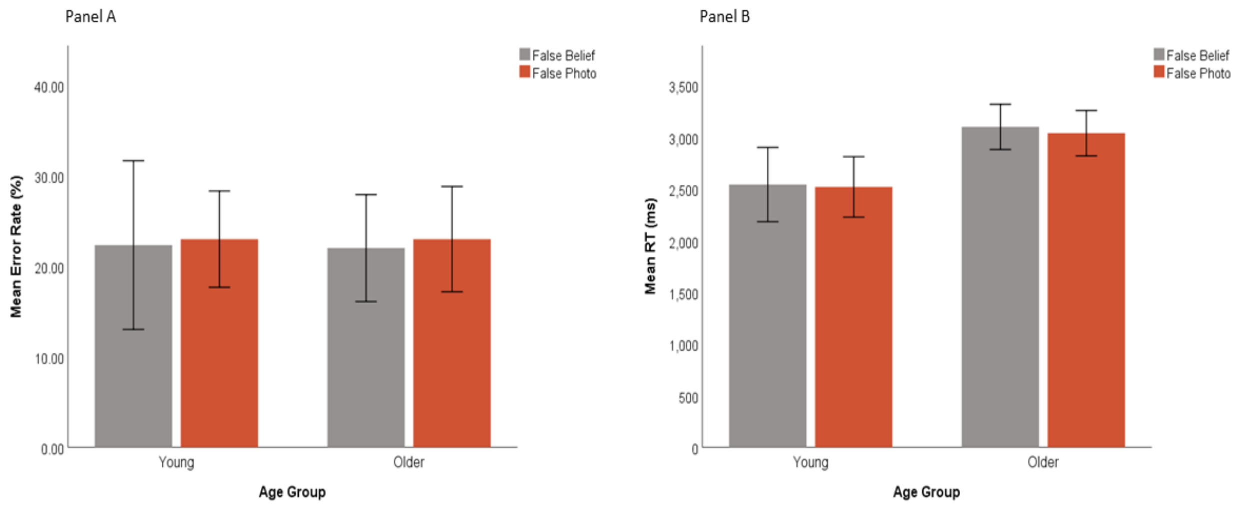


Figure 4.8. *Accuracy (error-rate, %; Panel A) and RT (Panel B) means as a function of Age Group across both experimental conditions. Error bars represent standard error (x 2). Grey bars represent False Belief condition and orange bars represent False Photo condition.*

First, with respect to accuracy data, there was no effect of Condition, $F(1, 36) = .097$, $p = .757$, $\eta p^2 = .003$; nor an interaction behind Condition and Age Group ($F(1, 36) = .004$, $p = .951$, $\eta p^2 = .000$). Finally, there was also no effect of age on accuracy, $F(1, 36) = .002$, $p = .968$, $\eta p^2 = .000$.

Next, in terms of response-time data, there was no effect of Condition on RT, $F(1, 36) = 1.052$, $p = .312$, $\eta p^2 = .028$. Likewise, there was no interaction of Condition on Age Group, $F(1, 36) = .217$, $p = .644$, $\eta p^2 = .006$. However, there was an overall effect of Age Group, $F(1, 36) = 7.808$, $p = .008$, $\eta p^2 = .178$. On average, OA ($M = 3070.924$, $SD = 478.608$) were ~540 ms slower compared to their younger counterparts ($M = 2531.866$, $SD = 710.400$).

Neuroimaging results

The analyses of brain data was largely modelled on the results of our behavioural findings. Namely, we examined potential contrasts between levels of each of the two factors in our current FB task. More precisely, to examine the neural consequence of

KOR, we modelled a reality unknown > reality known (i.e., KOR 0 > KOR 1) contrast, and for SOC, we modelled incongruence > congruence (SOC 1 > SOC 0). Further, as there was a KOR x Age interaction in the behavioural RT data, we also performed a series of additional contrasts to scrutinise any potential interactive effects on a neural level. These contrasts were then compared on a between-groups basis (young ≠ older) via a series of two-sample t-tests, using both GLM and non-parametric permutation-based models.

Knowledge of Reality

A comparison of the BOLD expression in KOR 0 > KOR 1, on a whole-brain level, revealed increased recruitment of the fusiform gyrus, bilateral TPJ, posterior cingulate gyrus, in addition to visual areas such as the calcarine sulcus. Overall, not having a knowledge of reality, where the location of the ball was unknown, resulted in greater activation in these areas compared to having an explicit knowledge of reality. See Figures 4.9 and 4.10 for illustrations of the results, and Table 4.12 for details.

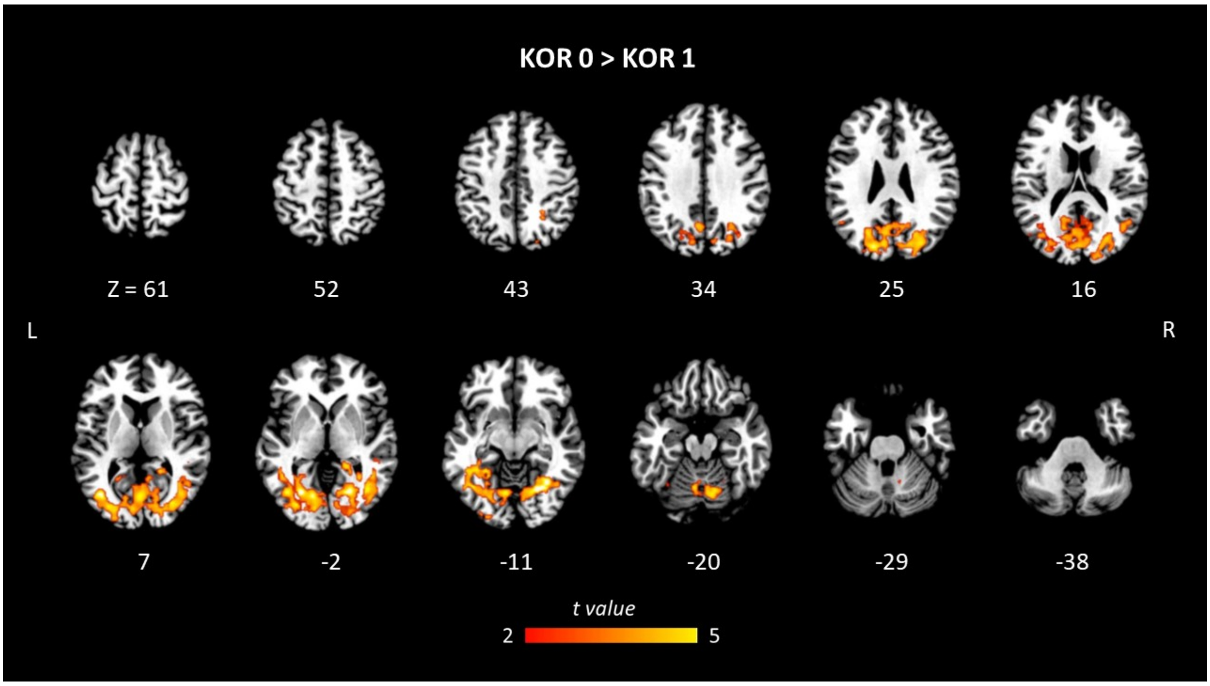


Figure 4.9. Selected axial slices illustrating significant clusters of differential BOLD response to the KOR 0 > KOR 1 contrast, after thresholding ($p< 0.05$) with non-parametric permutation bootstrapping. Parametric t-maps resulting from GLM analyses, thresholded by non-parametric permutations, are shown overlaid onto a common high-resolution brain template, within MNI space.

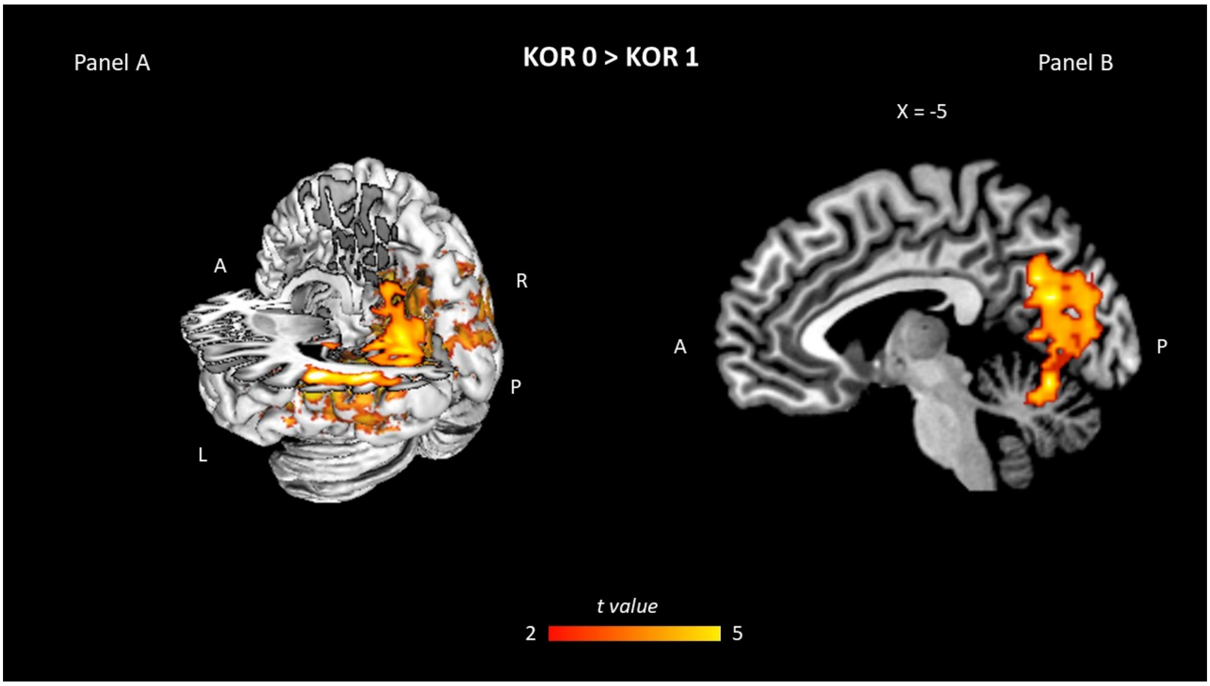


Figure 4.10. Panel A: Three-dimensional reconstruction of the BOLD expression in KOR 0 > KOR 1. Panel B: Single sagittal slice (X = -5, Y = 0, Z = 0) representing KOR 0 > KOR 1 contrast. Parametric t-maps resulting from GLM analyses, thresholded by non-parametric permutations ($p < .05$), are shown overlaid onto a common high-resolution brain template, within MNI space.

Table 4.12
Regions associated the KOR 0 > KOR 1 contrast

Cluster peak area	Cluster size (no. of voxels)	Max *Z*	Max X (mm)	Max Y (mm)	Max Z (mm)	COG X (mm)	COG Y (mm)	COG Z (mm)
Fusiform gyrus	18545	5.83	26	-66	-9	-0.356	-69.9	1.8
r TPJ*	1424	3.77	30	-80	25	26.6	-78	22.3
Cingulate gyrus (posterior)	372	4.12	19	-42	0	19.6	-43.9	1.26
Calcarine sulcus	109	3.55	-16	-74	9	-16.2	-74.5	11.5

Note: *TPJ was recruited significantly bilaterally, however, for clarity, only clusters with a $t > 3$ (thresholded at $p < .05$) are shown in the table. rTPJ = right temporoparietal junction. COG = centre of gravity.

We predicted that KOR would tap non-social executive (frontal) regions and also the anterior TPJ. We found that the KOR effect bilaterally recruited the TPJ, but the effect was strongest in the right hemisphere. To investigate the KOR-TPJ overlap more closely, the KOR effect was examined with Mars et al.'s (2012) connectivity-based, probabilistic parcellation. This parcellation subdivides the TPJ into anterior and posterior sub-regions, each region of the TPJ supporting attention and social-cognitive function, respectively. We found that the effect of KOR had a 62% overlap with the TPJa and only a 25% overlap with TPJp, in line with our expectations. In order to visualise the overlap, we segmented Mars' original mask into TPJa and TPJp composite masks and then registered these composites separately onto a standard MNI template brain space. These were then binarised individually to create discrete anterior and posterior sub-regions and added together in one single image. Figure 4.11 below illustrates the two discrete sub-regions, with TPJp coloured in yellow and TPJa in orange, and the overlap with the effect of KOR in white. It should be noted, however, the maps only show the right TPJ as Mars et al.'s parcellation only focused on the right hemisphere. This is complementary to our own analysis as our KOR effect was shown to be stronger in the right TPJ relative to the left

TPJ, where in fact KOR activation in the left TPJ did not yield a sufficiently large cluster to meet our *t* threshold (i.e., > 3).

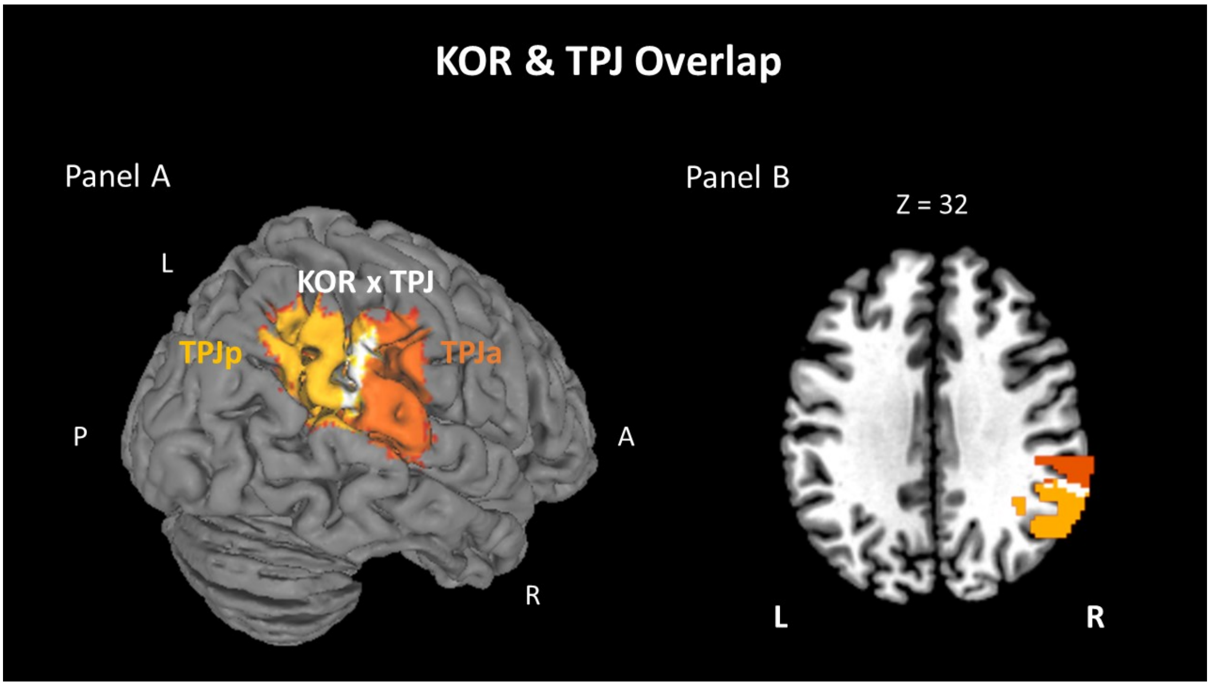


Figure 4.11. Illustration of overlap of KOR effect with posterior and anterior TPJ. Yellow indicates TPJp and orange indicates TPJa, and white highlights the KOR effect situated within the parcellated TPJ map. Mars et al.'s (2012) probabilistic map showed that there is a 62% overlap of the KOR effect with the TPJa, and a 25% overlap with TPJp. Panel A: Three-dimensional rendering of KOR-TPJ overlap, showing only the right hemisphere. Panel B: Axial view of KOR effect (white) visualised superimposed onto the anterior and posterior TPJ, orange and yellow, respectively.

Self-Other Congruence

As with our behavioural results, we did not find a significant effect associated with an incongruence > congruence contrast for our SOC factor, either with conventional GLM analyses or with non-parametric permutation testing. However, this effect (along with all other predicted FB effects) were scrutinised through an ROI approach – as informed by the ToM localizer (presented later in Chapter).

Self-Other Congruence x Knowledge of Reality interaction

To investigate any potential interactive effect of KOR and SOC, two further contrasts were run. First, we looked at K0S0 > K0S1 – where 'K' represents the KOR factor and 'S' represents Self-Other Congruence. The values '0' and '1' depict the theoretical presence/absence of interference in belief reasoning within that specific factor. As such, '0' refers to reality unknown and congruence, respectively, and '1' refers to reality known and incongruence, respectively. Thus, a K0S0 > K0S1 contrast specifically examines the effect or cost of SOC (congruence vs. incongruence) when reality is not known to the participant (i.e., KOR 0).

A whole-brain analysis of the K0S0 > K0S1 contrast, after non-parametric permutation thresholding at $p < .05$, revealed significant BOLD expression primarily in the precuneus and left TPJ, and to a lesser extent, the right STS and anterior cerebellum (vermis). See Figures 4.12 and 4.13 for illustration of results, and Table 4.13 for details.

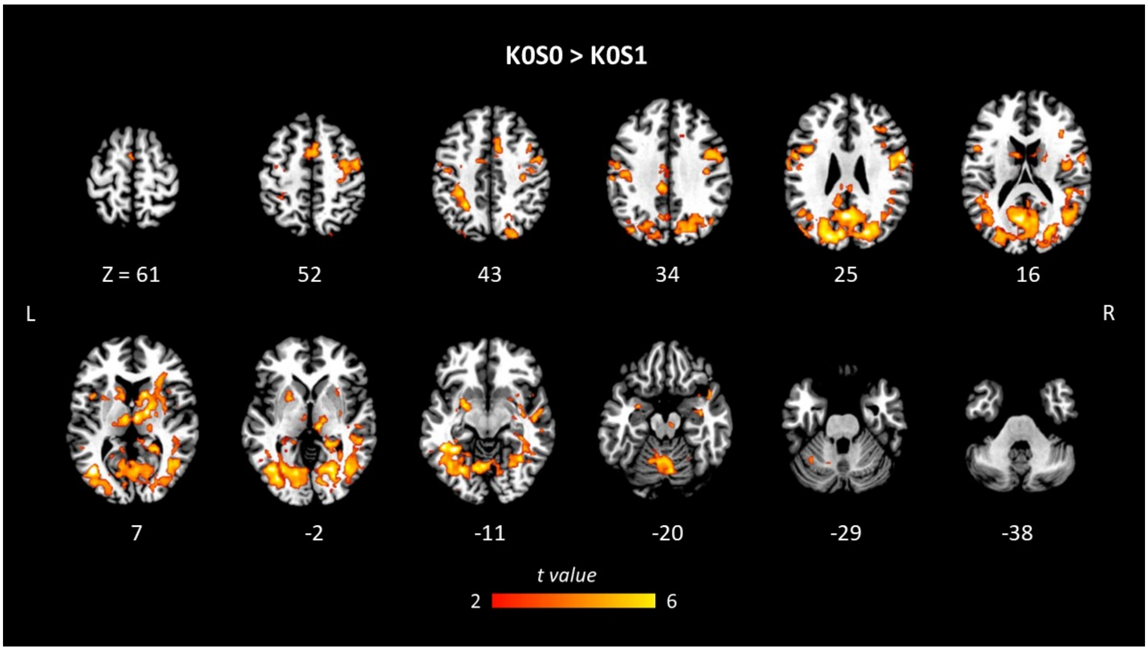


Figure 4.12. Selected axial slices illustrating significant clusters of differential BOLD response to the K0S0 > K0S1 contrast. Parametric t-maps resulting from GLM analyses, thresholded by non-parametric permutations ($p < .05$), are shown overlaid onto a common high-resolution brain template, within MNI space.

Table 4.13.
Regions associated with the K0S0 > K0S1 contrast

Cluster peak area	Cluster size (no. of voxels)	Max *Z*	Max X (mm)	Max Y (mm)	Max Z (mm)	COG X (mm)	COG Y (mm)	COG Z (mm)
Precuneus	12829	4.87	-2	-64	21	5.81	-69.8	13.9
L TPJ	6284	5.49	-41	-70	5	-32.2	-68.4	-2.83
R STS	2385	4.38	20	-79	25	24.1	-77.3	24.5
Cerebellum (anterior lobe)	1267	4.06	-4	-64	-11	-3.65	-65.6	-15.1

LTPJ = left temporoparietal junction. COG = centre of gravity. RSTS = right superior temporal sulcus.

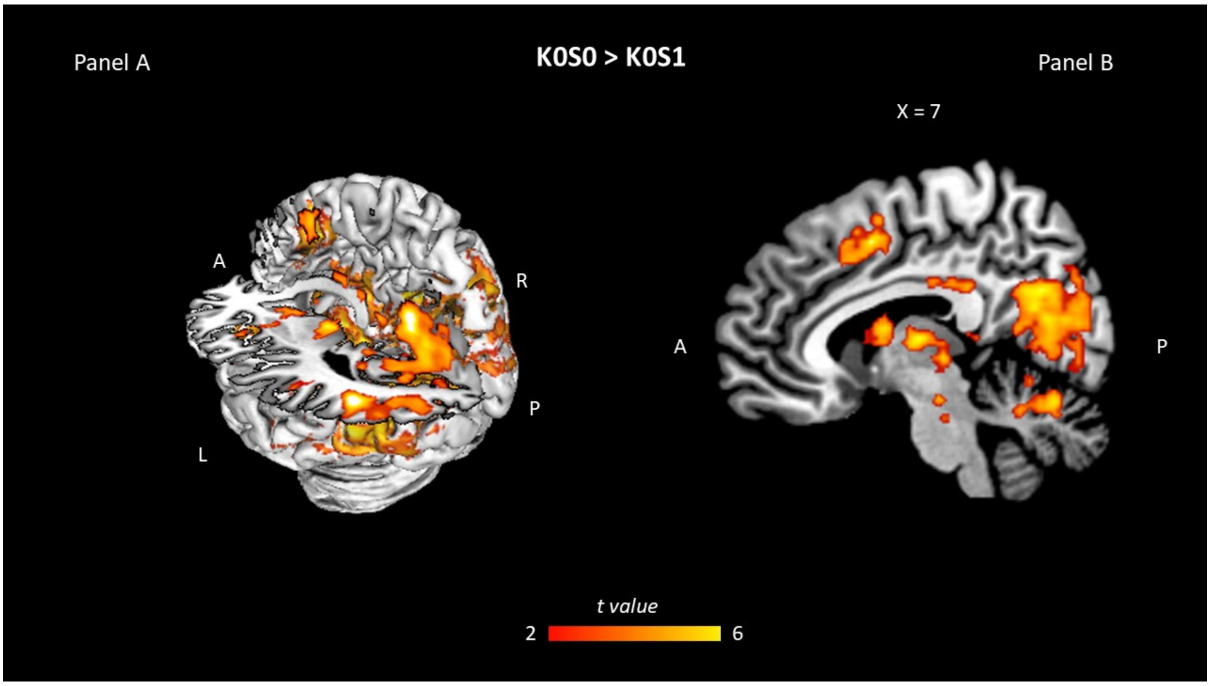


Figure 4.13. Note: Panel A: Three-dimensional reconstruction of the BOLD expression in K0S0 > K0S1. Panel B: Single sagittal slice (X = 7, Y = 0, Z = 0) representing K0S0 > K0S1 contrast. Parametric t-maps resulting from GLM analyses, thresholded by non-parametric permutations ($p < .05$), are shown overlaid onto a common high-resolution brain template, within MNI space.

A second contrast was performed to delineate further the interaction between KOR and SOC: K1S1 > K1S0. Here, we tested the cost of self-other congruence (congruent vs. incongruent) when participants had an explicit knowledge of reality (i.e., KOR 1). This condition is analogous to a classic false vs. true belief effect, in that the participant is explicitly aware of reality and must represent a belief state which is either congruous (true belief) or incongruous (false belief) with the agent's.

No significant regions or clusters were identified through this contrast – either through GLM or permutation-based models – and therefore we have not reported these here. However, as previously mentioned, all contrasts were modelled and are reported in our ROI analyses (later in chapter). We likewise did not see any age-related effects based on our proposed contrasts; though, these are explored through ROI analyses later in the chapter.

ToM localizer

Briefly, as previously discussed, the theory of mind localizer task (Saxe & Kanwisher, 2003) comprises two conditions: Belief and Photo. While the Belief condition requires participants to mentalize and reason about others' mental states, the Photo condition can be resolved through non-social inferences. Therefore, in line with Saxe and Kanwisher (2003), we modelled a Belief > Photo ('B > P' henceforth) contrast to delineate between these processes and expected to see greater recruitment of areas within the classic *mentalizing network* (Abu-Akel and Shamay-Tsoory, 2011). More importantly, regions highlighted within this analysis were used to inform a subsequent ROI analysis of the proposed effects in our FB task.

First, the B > P contrast showed that the Belief condition – collapsed across both age groups – led to greater recruitment of bilateral TPJ, medial prefrontal cortex (mPFC), precuneus, inferior occipital gyrus, and the left DLPFC. The regions highlighted here are broadly similar to that of the original authors (Saxe & Kanwisher, 2003; Dodell-Feder et al., 2011). See Figures 4.14 and 4.15 for illustration of results, and Table 4.14 for details.

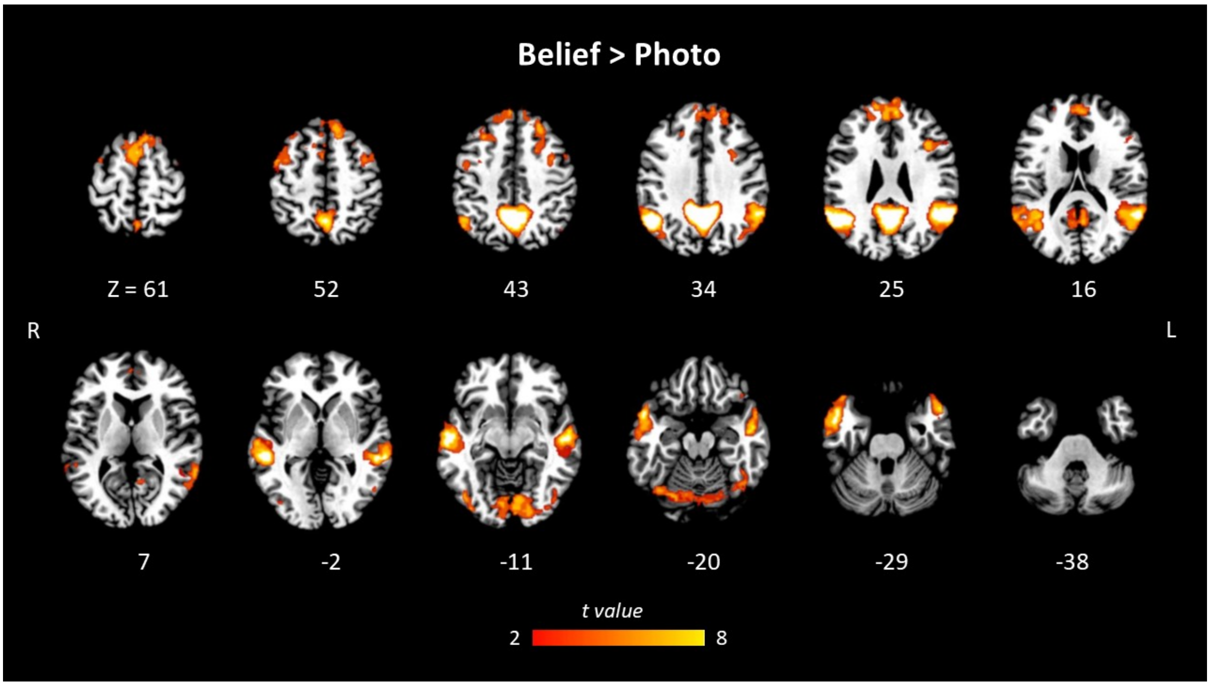


Figure 4.14. Axial layout view highlighting the significant clusters associated with a Belief > Photo. Parametric t-maps resulting from GLM analyses, thresholded by non-parametric permutations ($p < .05$), are shown overlaid onto a common high-resolution brain template, within MNI space.

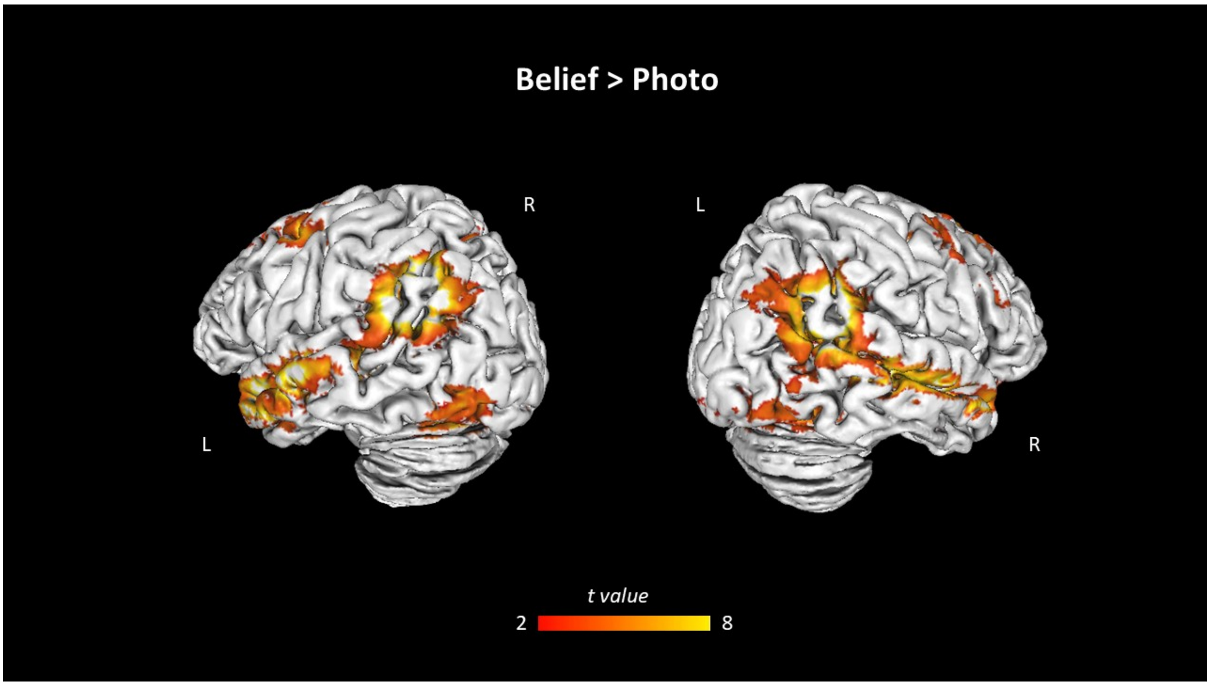


Figure 4.15. Left/right three-dimensional rendering of the expression of Belief > Photo contrast from ToM localizer task. Note the particular recruitment of both the left and right TPJ. Parametric t-maps resulting from GLM analyses, thresholded by non-parametric permutations ($p < .05$), are shown overlaid onto a common high-resolution brain template, within MNI space.

Table 4.14.
Regions associated with the Belief > Photo contrast

Cluster peak area	Cluster size	MAX Z	MAX X (mm)	MAX Y (mm)	MAX Z (mm)	COG X (mm)	COG Y (mm)	COG Z (mm)
L TPJ	4426	7.07	-52	-60	30	-52.7	-35.6	7.4
mPFC	4374	4.99	-4	12	68	9.49	30.5	43
R TPJ	3859	6.1	60	-50	28	54	-37.3	10.1
Precuneus	3523	7.26	-6	-52	42	0.19	-55.2	36.2
InferiorOcc	2029	4.94	16	-86	-16	-1.05	-78.1	-13.3
L DLPFC	706	4.77	-50	0	52	-37	10.8	48.5

Note: L/R = Left, Right. TPJ = Temporoparietal junction. mPFC = Medial prefrontal cortex. InferiorOcc = Inferior occipital gyrus. DLPFC = Dorsolateral prefrontal cortex.

When comparing young vs. older adults, we did not find significant differences in activations maps in the ToM localizer data.

ROI analyses

Next, we scrutinised all four of our a priori hypotheses/contrasts from the FB task – and their differentiation as a function of age – within each node of the mentalizing network engaged during the localizer task. Spheres of 10 mm radius were created centred on the voxel which maximally expressed the B > P contrast within each highlighted region (see Table 4.14). This led to the creation of six spheres: left TPJ, right TPJ, precuneus, mPFC, inferior occipital gyrus, and the left DLPFC. Within each of these spheres/clusters, we examined our four original a priori contrasts between the two age groups: 1) KOR 0 > KOR 1, 2) SOC 1 > SOC 0, 3) K0S0 > K0S1, and 4) K1S1 > K1S0. The first two contrasts represent the main effects of Knowledge Of Reality and Self-Other Congruence, respectively, whereas the latter two contrasts reflect interactions behind these two effects. The last contrast, K1S1 > K1S0, as discussed, is a classic false versus true belief contrast in that it compares, when reality is known, incongruence (false belief) vs. congruence (true belief) between self and other.

This ROI approach was useful here to examine the extent to which there was a differentiation between age groups in the expression of these contrasts within circumscribed, localizer-informed regions of the brain. While we previously did not find age-related differences on a whole-brain level – we suspect, largely due to low statistical power – scrutiny of these select regions allowed us to examine age effects within established nodes of the mentalizing network. To this end, we examined mean COPE (contrast of parameter estimates) values of each contrast within the predefined regions, as a function of age.

Four separate 2 (Age: YA vs. OA) X 6 (Regions: L TPJ, mPFC, R TPJ, Precuneus, Inferior Occ, L DLPFC) mixed ANOVAs (one for each contrast) were run to examine between-group differences in activation maps within localizer-defined regions – see Figure 4.16.

Mean parameter estimates of each FB contrast within localiser-informed ROIs, as a function of age.

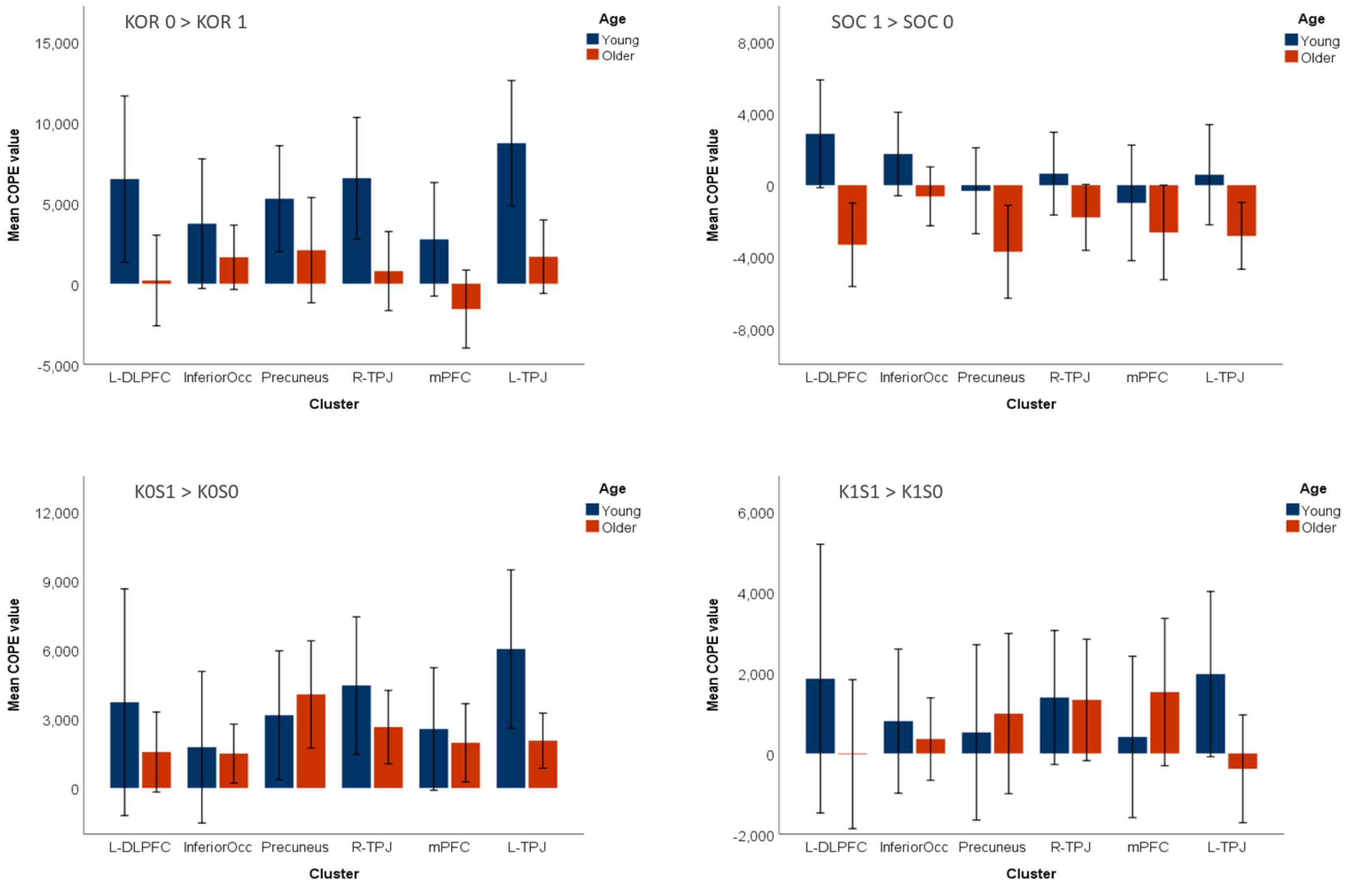


Figure 4.16. Mean parameter estimates of each FB contrast within localizer-informed ROIs, as a function of age. KOR 0 > KOR 1: reality unknown > reality known. SOC 1 > SOC 0: self-other incongruent > self-other congruent. K0S1 > K0S0 (interaction effect): reality unknown, self-other incongruent > reality unknown, self-other congruent. K1S1 > K1S0 (interaction effect): reality known, self-other incongruent > reality known, self-other congruent. Navy bars represent younger adults and orange bars represent older adults.

The analysis did not reveal significant between-group differences in any of the four contrasts.

As can be seen from Figure 4.16, scrutiny of each contrast on a region-by-region basis as a function of age highlights emerging (though non-significant) differential recruitment of certain cortical areas. In terms of interpretation of the mean COPE values, a positive (greater than zero) value indicates an effect in the direction consistent with the proposed contrast. In the KOR 0 > KOR 1 contrast, therefore, younger adults exhibited (emerging) greater recruitment of the left and right TPJ, left DLPFC, and mPFC.

While the cost of incongruence in self-other beliefs (SOC 1 > SOC 0) did not emerge as a main effect in the whole-brain or ROI analyses, there was noticeable qualitative differentiation between the age groups when examining the contrast within the predefined ROIs. The difference in regional activation was particularly evident in the left DLPFC and precuneus – though in the left DLPFC, while the effect for YA was directionally consistent with the proposed contrast, the effect for OA opposed the planned contrast. For the two interactions, there was some evidence of differentiation between age groups, particularly in the left TPJ.

It should be noted, however, these differentiations were not substantiated statistically. Linked to this, due to the previously-described disruption to our study protocol, our current design does not possess adequate power to detect such differences. However, these data nevertheless highlight emerging patterns of differential recruitment of circumscribed brain regions between YA and OA groups that will be investigated fully once the issue of power can be fully addressed.

Discussion

In this study, using fMRI, we set out to delineate known sources of interference in belief reasoning and examine if these sources of interference differentially affected performance as a function of age. Unlike prior fMRI studies of ToM, our work includes an important adjustment within our fMRI analysis for individual differences in neurovascular integrity, which is particularly relevant when considering age (Tsvetanov et al., 2015; 2021). Unfortunately, we were somewhat curtailed due to COVID-19-related disruption. This affected planned participant recruitment and subsequent data analysis. Thus, this work should be considered as preliminary.

On a behavioural level, we were broadly able to replicate our original findings – especially with respect to our experimental manipulation of *knowledge of reality;* all participants were slower to respond when the location of the object was unknown. We did not, however, replicate our findings with regards to the effect of SOC. Previously, we had found that all participants were slower and more error-prone when responding to an agent with an incongruent belief. The present data did not show these effects, although the direction of findings were consistent with this. On a neural level, the localizer task recruited similar regions as to what have been previously published. We showed differential brain responses based on our theorised, relevant sources of social and non-social interference. However, we were unable to show age effects in the neuroimaging data. Each of these effects and their implications are discussed separately below.

Knowledge of Reality

We experimentally manipulated the presence/absence of a knowledge of reality in our novel FB paradigm and found that participants were faster when they were explicitly made aware of the true state of affairs. However, classic literature on the sources of interference in FB reasoning has suggested that due to an egocentric bias or a 'curse of knowledge' (see Birch and Bloom, 2003, 2007), a prominent knowledge of reality compromises one's ability to reason about other people's false beliefs. Birch and Bloom (2007) reported that, from their study using a Sally-Anne-like displacement task, adults who had privileged knowledge of the specific outcome of a displacement event were better able to reconcile a FB problem, compared to adults who were naïve. Similarly, Mitchell and colleagues (1996) note that a privileged knowledge of the truth 'contaminates' belief processing in adults by a phenomenon known as *realist bias*. Contrary to this, we found that reasoning about a FB is less difficult when knowledge of the true state of affairs is revealed; in other words, not knowing reality led to poorer performance. This finding is a replication our previous result in Chapter 3. However, the effect of knowing reality evidenced in the current study and the classic *curse of knowledge* hypothesis are not necessarily in discord when understood in the context of belief-desire reasoning, based on the target agent's intentions (see Friedman & Leslie, 2004; Leslie & German, 2005). The interference from content knowledge may not be from the actual unveiling of privileged information, but perhaps based on additional processing that is required to manage third-party intentionality. More precisely, the additional interference from the reality unknown condition could reflect an avoidance-desire mechanism, whereby the participant additionally processes the agent's implied desire to avoid an empty location. This interpretation is consistent with Leslie's 'double inhibition' model (Friedman & Leslie, 2005; Leslie et al., 2004; Leslie et al., 2005) where participants process two inhibitions to successfully resolve the task: first, the participant inhibits their privileged knowledge of reality (to represent a false belief), and second, the participant processes and inhibits the empty location in line with the agent's desire or implied intention to avoid the vacant site. This additional processing could, at least in part, explain faster responses in the reality known condition (compared to reality unknown). In line with this interpretation, studies examining the time-course of representing others' perspectives have shown that there is a bias toward seeking to fulfil, not conflict, the desires of others, where participants may anticipate intention or preference-compatible – in the present case, the agent seeking the ball – outcomes (e.g., see Ferguson & Breheny, 2011).

On a neural level, for the effect of KOR, we expected to see significant clusters of activation within regions classically associated with EF (i.e., frontal areas) in addition to areas involved in attention, including a specific region of the TPJ. Studies investigating the structural and functional connectivity of the TPJ have shown that the area can be subdivided into discrete sub-regions, serving both attentional (anterior) and social cognitive (posterior TPJ) functions (see reviews by Krall et al., 2015; Mars et al., 2012). In the present study, as our KOR factor was theorised to rely on wider domain-general executive (non-social) processes, we expected to see recruitment of the anterior TPJ when resolving reality known vs. reality unknown conditions due to its association with attentional control (see Geng & Vossel, 2013 for a review on the TPJ and its role in attention). Accordingly, and as expected, we found greater activation of the anterior TPJ supporting our hypothesis that managing conflict from a knowledge of reality manipulation within the context of ToM reasoning is perhaps more reliant on non-social, cognitive mechanisms such as attention or attentional control. This result indicates that activation of the TPJ in fMRI studies involving FB comparisons may not entirely reflect social cognitive processes. Therefore, activation of the TPJ cannot be used to infer that a task is tapping social processes without careful mapping of this activation. Indeed, the term ‘TPJ’ is used to describe a relatively large area of the cortex; however, as outlined earlier in this chapter, there is growing evidence of functional differentiation within this region. Thus, future work should ensure that studies are carefully designed to ensure focus is on the relevant neuroanatomical region of interest. This could be achieved through carefully designed functional, within-subject comparisons or through the use of a localizer task, for example.

Overall, we found a behavioural effect of KOR where responses were faster in the reality known condition, a finding that we initial observed in Chapter 3. However, this main effect was driven largely by the OA group, where an Age x KOR interaction showed that there was only a significant reduction in response-times in reality known vs. unknown in the older group. In the fMRI data we found significant activation, among other regions, in the TPJ and we were able to show specificity of this effect particularly within the anterior TPJ, in line with expectations. With a fully powered study, we would be able to look with greater confidence as to whether differential recruitment of the anterior TPJ might explain these differences.

Self-Other Congruence

We expected to see differential behavioural performance according to congruence in Self-Other beliefs. By manipulating congruence, we were able to test specifically a true vs. false belief condition, where, in line with previous literature, we expected greater interference from the Self-Other incongruent condition. In comparison to congruent Self-Other (true) beliefs, we hypothesised that responses would be slower and less accurate when participants resolve Self-Other incongruent (false) beliefs, and that this cost would be proportionately greater in OA. Prior work has shown that inhibiting the self-perspective, which is a precondition of representing and reasoning about false beliefs, introduces additional processing and is dissociable from wider EFs (e.g., Hartwright et al., 2015; Samson, Apperly, Kathirgamanathan, & Humphreys, 2005); and has been shown to pose particular difficulty in advanced age (e.g., German and Hehman, 2006). Further, we hypothesised that SOC would drive significant recruitment of regions or nodes within the classic mentalizing network, chief among which is the (posterior) TPJ. In line with this, Schurz et al. (2014) identified the medial PFC and bilateral posterior TPJ as core networks when reasoning about false beliefs, irrespective of task and stimulus formats. However, we did not find an effect of SOC in the behavioural or functional neuroimaging analysis. This contrast was also further scrutinised in a follow-up ROI analysis. Again, we should be cautious until the full study is completed. However, given that prior research suggests that cross-subject variation in fronto-parietal networks increases with age (see Li et al., 2017), it is probable that the lack of findings reflect lack of power. This might seem counterintuitive, given that we were able to identify an effect for KOR; however, it is possible that there is greater functional variability in how domain-general processes, like attention, are managed when compared with social processes, like ToM. Moreover, it is also not possible to determine at this point whether prior findings are affected by age-related differences in neurovasculature, which we control for in our analysis of the novel ToM task (but do not in the localizer to permit wider comparability). Still, what is notable from the ROI analysis (which was derived from the localizer) is that there were qualitative differences between age groups in how the brain responded when managing conflicting self-other perspectives. We will assess this quantitatively once we are able to finalise the research.

Knowledge of Reality, Self-Other Congruence, and the effect of Age

We previously reported a KOR x SOC interaction in our behavioural FB investigation (Chapter 3), where, when reality was unknown (KOR 0), responses were more accurate in the SOC congruent (SOC 0) relative to SOC incongruent (SOC 1). There was previously also an effect of KOR in the SOC congruent condition, where participants responded more accurately in the reality unknown (KOR 0) condition relative to the reality known (KOR 1) condition. With the current data, however, there was no interaction behind the KOR and SOC manipulations, either in the RT or accuracy behavioural data.

On a neural level, we found a significant interaction between KOR and SOC where in reality unknown condition an incongruence in self-other beliefs highlighted several regions within the core mentalizing network, namely, the left TPJ in addition to the precuneus and right STS.

Behaviourally, we observed an interaction in the effect of knowledge of reality with age. Overall, a privileged knowledge of reality reduced response-times, as noted above, however, this benefit (of knowing vs. not knowing reality) was only observed in OA – the KOR manipulation had no effect on YA. This finding diverges from our initial behavioural study (Chapter 3) as previously we found no interaction with the KOR manipulation with age.

While we did not find statistically-relevant age effects in the functional data, it is worth considering the emerging group differences in the ROI analyses. We employed Saxe and Kanwisher's (2003) ToM functional localizer to highlight regions of interest, within which we focused our investigation of age differences in the context of our FB paradigm. We observed incipient age-related effects, across all four of our planned contrasts, in regions relevant to the mentalizing network (see Abu-Akel & Shamay-Tsoory, 2011). For example, for the contrast of reality unknown > reality known (KOR 0 > KOR 1), there was differentiation within the TPJ (bilateral), left DLPFC, and mPFC. While age-related differences in the recruitment of these areas did not survive statistical thresholding, it is plausible that with sufficient statistical power such differences would reliably show.

Further, the incongruent > congruent contrast in self-other beliefs did not produce statistically relevant neurobehavioural results, as a standalone main effect or an interaction with age. However, an inspection of the mean contrast parameters shows that there was noticeable between-groups differentiation in the recruitment of localizer-informed regions, chiefly within bilateral TPJ and left DLPFC, regions that have been implicated extensively in the prior belief-reasoning literature. To our knowledge, the functional neuroimaging literature is yet to show a self-other beliefs distinction in the context of ageing. However, there is an emerging literature base which highlights key regions – particularly the TPJ – in the distinction between Self and Other in a range of different socio-cognitive operations, including mental state representation, social imitation/mimicry, and empathy (Quesque & Brass, 2019).

Contribution of current findings to wider cognitive ageing literature

The current work complements and extends the literature on general cognitive ageing. As has been shown in the literature, we report poorer performance in OA in abilities related to memory, processing speed, attention, and inhibition, and found that these measures, to varying degrees, are associated with age-related decline in ToM performance. From both the quantitative review and behavioural investigation, we showed that age-related decline in domain-general functions is highly variable, with higher EF capacities predicting superior ToM performance in older age. Further, in our neuroimaging investigation, we again found that OA were more varied in both their behavioural and neural response in comparison to their younger counterparts. Increased variability in advanced age is thought to derive, in part, from modifiable lifestyle factors including education, physical/cardiovascular health, individual differences in frailty and mobility, size of social networks, socio-economic status, and more (Tucker-Drob & Salthouse, 2011). However, it is difficult to draw a definitive, causal account of what leads to ToM decline in old age, though it is clear that declining EF plays a significant role. No investigations to date have factored in and explored how background lifestyle and environmental factors interact with cognitive ageing in its relationship with SCF. The cognitive ageing and ToM literature would benefit from such a study, where flexible social factors (such as SES, education, social participation, etc.) are considered when delineating the link between diminishing cognition and ageing ToM.

Limitations and future directions

There are limitations that should be considered in the context of this study. First, while the study employed a measure of reading ability through the WRAT-3, the current design did not include a measure of executive function. Previous research has shown that age-related decay of socio-cognitive aptitude may in part be explained by changes in domain-general abilities (e.g., Rakoczy, Harder-Kasten, & Sturm, 2012). Indeed, in our own work (Chapter 3) we have shown that the cognitive mechanisms underlying FB reasoning may in part be explained by variability in attention switching and inhibitory control. That being said, it should be noted that, at the time of writing, there is a significant dearth in the functional imaging literature associating executive function with age-related changes in ToM. What is more, participants' highest level of education – which has been widely linked with variability in EF – was roughly equivalent between the two cohorts. Overall, to further clarify the role of EF in ageing ToM in the context of neuroimaging designs, future investigations should encompass EF measures (in particular, inhibition) to delineate on a neural level the contribution of cognitive control when resolving problems of false belief.

Another limitation of the current study is the sex imbalance in the overall sample, which was due to COVID restrictions preventing full data collection. While the male:female proportion was equivalent across both age groups, within each group the sample was still roughly 75% female. An imbalance in males and females in the sample is potentially a noteworthy concern as previous studies have found sex differences in adult mindreading abilities (e.g., see Krach et al., 2009). Therefore, it is feasible that an omission of sex from the current design may have veiled underlying behavioural and neural effects and that the addition of sex as a factor could reveal meaningful interactions with the other experimental manipulations in our task. However, in a supererogatory analysis where sex was included as a factor, there was no interaction with either KOR or SOC and/or Age (all *p*'s > .05). While this shows that sex was perhaps not a statistically influential factor in our results, it is again worth caveating this discussion with the issue of low statistical power due to our curtailed sample size.

In the current study, the inclusion criteria with respect to age was 18-29 for young participants and 60-79 for older participants. With respect to our older cohort, the difference between the youngest and oldest participant (62 and 77, respectively) was 15 years. We did not further differentiate on the basis of age within the older adult group, despite the seemingly large age range. Indeed, previous investigations of ageing ToM have commonly demarcated older participants into young-old and old-old categories (typically, 60-70 and 70>, respectively) and have reported behavioural group differences in mentalizing abilities (e.g., Baksh et al., 2018; refer to Chapter 3 for a review of ageing ToM studies). Thus, by grouping all older adults into a single category we perhaps may have masked variability in both ToM performance and the subsequent neural consequence in this group. It is worth noting, however, low statistical power precluded any sub-group analyses of our older adult data. Thus, for a more granular neurobehavioural account of socio-cognitive abilities in ageing populations, and how these abilities may vary *within* ageing, future ToM neuroimaging investigations should separate participants into smaller, age-constrained categories.

Conclusion

Through our novel false belief paradigm, we tested young and older adults to delineate the neurocognitive mechanisms involved in reasoning about others' beliefs. On both a behavioural and neural level, we differentiated between possessing (vs. not possessing) a privileged knowledge of reality and found this to improve performance, and highlighted distinct (and, in some cases, unexpected) cortical regions – namely, the right TPJ and fusiform and cingulate gyri – involved in this process. We were also able to show specificity of the knowledge of reality effect within the anterior TPJ, a region thought to be associated with wider attention mechanisms. We further observed an interaction between knowledge of reality and self-other congruence, again both on a behavioural and neural level, with classic mentalizing nodes emerging as significant clusters, including the precuneus, TPJ, and STS involved in resolving incongruence in self-other beliefs when participants did not possess a privileged of reality. However, we did not observe differences based on a core manipulation of self-other congruence ([in]compatibility in self-other beliefs), either in the behavioural or neuroimaging data. While we showed emerging group differences in certain cortical regions, as informed by a ToM localizer, we did not observe statistically relevant effects or contrasts between young and old in the neuroimaging data. Thus, this work provides a foundation for understanding the mechanisms involved in representing others' mental states and the neural consequence of each of these effects, and provides a localized basis on which age differences in these effects may occur at a neural level.

Chapter Five: General Discussion

Two broad aims were the focus of this thesis. First, the nature and magnitude of age-related differences in ToM reasoning were investigated, with the aim of conveying greater clarity and specificity with respect to the classically reported effect of age on socio-cognitive function. Next, the process of reasoning about others' mental states was decomposed to investigate the precise neurocognitive sources of difficulty and interference when engaging ToM, and how these sources of interference may differentially affect young and older healthy adults. The first aim was addressed primarily through a meta-analysis of behavioural ToM-ageing data in Chapter 2. The second aim was addressed through two empirical investigations in Chapters 3 and 4. Chapter 3 sought to disentangle processes thought to be important in the context of FB reasoning, which are often confounded in the literature. Further, the Chapter highlighted social and non-social cognitive processes which differentially affect the performance of older adults relative their younger counterparts. Chapter 4 extends the findings of Chapter 3 through an fMRI study using a modified version of our novel FB paradigm. After correcting for age-related differences in neurovascular coupling, which in itself is a novel addition to the literature, we explored the neurocognitive bases that underlie FB processing in ageing through both whole-brain and localised approaches.

In this General Discussion I will discuss the wider theoretical and methodological implications of the findings presented in this thesis, in addition to highlighting limitations and future directions, before ending with a Conclusion. First, however, I will summarise briefly the rationale and core findings of each Chapter.

Chapter Summaries

Chapter 2. There is substantial heterogeneity in the literature in both the theoretical account of what ToM actually constitutes and in the design and implementation of tasks and protocols. This non-standardised approach has led to qualitative differences in task difficulty, which may at least in part explain the mixed nature of an age effect when comparing young vs. older adults. Tasks that pose more difficulty by way of introducing supplementary (non-social) cognitive demands, may amplify an effect of age not necessarily because of poorer socio-cognitive function in OA but because fronto-executive abilities fade in advanced age. Previous meta-analyses have found that the magnitude of an age effect varies depending on task and sampling considerations, and significant statistical heterogeneity exists among the ageing-ToM data, even when looking at paradigms that are purportedly similar in design. To address this issue, the nature and magnitude of an age effect was further examined through a meta-analysis of behavioural ToM-ageing data. In addition to examining an overall effect of age, we performed a series of theoretically-driven sub-group analyses to highlight the conceptual and methodological considerations that drive variability in ageing ToM. Specifically, we examined the effect of age as a function of task modality and domain and the demographic make-up of OA samples – i.e., age and education.

While a strong effect of age emerged overall, there was again significant heterogeneity in effects. In a small number of sub-group analyses, the effect of age was abolished completely, suggesting that the magnitude (and indeed presence) of age differences is at least in part contingent on methodological factors. Namely, we reported that the effect of age varied as a function of task characteristics: task domain (verbal vs. non-verbal) and task type (movies, static cartoon-comic, immersive VR, and vignettes). The difference in young vs. older ToM was more pronounced in non-verbal tasks and the greatest age difference was seen in the movies paradigms. The effect of age also varied according to sampling characteristics where greater age differences were seen in studies where OA were less well-educated. Also, as expected, a greater difference was shown when OA were greater in age – i.e., ≥ 75. Overall, in our review of the evidence, we found that ToM competence may indeed diminish with age, but importantly, the magnitude of this decline is modified – and in some cases, abolished – depending on task and sampling considerations.

Chapter 3. In the meta-analysis in Chapter 2, we showed that the effect of age is highly heterogeneous, varying according to both task and sampling considerations. One particular reason for this heterogeneity is perhaps down to the qualitative differences in task difficulty, where some ToM paradigms have greater non-social cognitive loads than others. In typical FB tasks, successful performance relies on both social and non-social (or executive, domain-general) processing. Non-social demands rely on intact EFs, which are also known to decline with age. However, within the ageing FB literature, this is yet to be disentangled; processing that is non-social in nature is often confounded with social cognitive function. We addressed this by testing how older versus younger healthy adults were affected by three theoretically relevant sources of conflict within ToM: competing Self-Other perspectives; competing cued locations, and outcome knowledge. We examined which best accounted for age-related difficulty with ToM. We also collected a battery of EF measures, including inhibitory control, attentional shifting, working memory, and processing speed. Our data showed similarity between age groups when representing a belief incongruent with one's own. Performance in attention and response speed tests best explained the degree of conflict experienced through incompatible Self-Other perspectives. However, OAs were disproportionately affected by managing conflict between cued locations. Age and spatial working memory were most relevant for predicting the magnitude of conflict elicited by conflicting cued locations.

Overall, FB reasoning was found to be effortful for OA beyond the non-social cognitive demands of classic ToM investigations. Performance in each of the ToM scenarios was related to individual differences in inhibitory control and spatial working memory. However, the magnitude of conflict experienced and the cognitive systems recruited to resolve this were condition specific: managing competing cognitive perspectives was supported by attentional resources whereas invalid cueing seemed to rely on spatial working memory. Further, OA were largely disadvantaged by invalid cueing. Altogether, our findings suggest that prior studies may have inflated the effects of ageing on ToM by including excessive conflict in ToM tasks.

Chapter 4. In Chapter 3, we investigated the sources of interference in FB reasoning and how they differentially affect older and younger participants. In Chapter 4, we extended these findings and studied the neurocognitive profile of ageing ToM by administering fMRI to both young and older adult participants while they responded to a revised version of our false belief paradigm and an established ToM functional localizer. Unlike prior fMRI studies of social cognition in ageing, our work includes an important adjustment of the fMRI data to account for individual differences in neurovascular integrity, which is particularly relevant when considering age differences.

On a behavioural level, we were broadly able to replicate our original findings in Chapter 3, particularly with respect to our experimental manipulation of KOR (or outcome knowledge) where participants were slower to respond when the location of the object was unknown. We did not, however, replicate our findings with regards to the effect of SOC. Previously, we had found that both groups of participants were slower and more error-prone when responding to an agent with an incongruent belief. Our data did not show these effects, although the direction of findings were consistent with this. On a neural level, the localizer task recruited similar regions as to what have been previously published. We showed differential brain responses based on our theorised, relevant sources of social and non-social interference. However, while we showed emerging group differences in regions informed by the ToM localizer, we were unable to show statistically significant age effects in the neuroimaging data.

It is important to note that the results of this study should be interpreted with caution due to low statistical power, especially with respect to between-group comparisons, and that no firm conclusions can be drawn based on these findings. Participant recruitment and testing was substantially curtailed due to COVID-19 social restrictions. At the time of writing, participant testing has resumed and once data collection has concluded, these data will be re-analysed in line with our original research questions and written up for publication. Therefore, while the core research questions and predictions will remain unchanged, it is possible that the findings presented within this Chapter may not be entirely consistent with our eventual write-up upon completion of the fMRI study.

Implications

Defining theory of mind

Although there is a general acceptance that ToM refers to a specialised set of socio-cognitive skills that facilitates the representation of others' mental states, the concept has been formally ill-defined. There are a multitude of different sub-components and processes that are characterised as ToM, even when such concepts ostensibly fail to satisfy the basic tenets of ToM functioning. Quesque and Rossetti (2020) argue that both vocabulary and measures lack specificity. For example, terms such as mentalizing, mindreading, perspective-taking, (cognitive) empathy, and mental state attribution, among others, all loosely refer to abilities commonly understood to mean *theory of mind*. The heterogeneity in terminology perhaps in part reflects the often confused conceptualisation, where *cognitive* and *affective* abilities are confounded. For instance, in its original publication, the Reading the Mind in the Eyes Test (RMET; Baron-Cohen et al., 1997) is referred to as an "advanced measure of mentalizing", and even though stimuli presented in the task require the attribution and recognition of both cognitive and affective states, the authors fail to distinguish between these processes. Baron-Cohen and colleagues simply entangle both processes as *mental state attribution*. Though commonly used as a measure of individual differences in ToM, a recent psychometric assessment of the RMET demonstrated that this task better reflects emotion processing, rather than mentalizing per se (see Kittel, Olderbak, & Wilhelm, 2021; Oakley et al., 2016). In addition to non-specific vocabulary, there is a considerable number of different measures and protocols in the literature. In their systematic review of ToM measures for young children, Beaudoin et al. (2020) assessed 830 studies that included 220 measures of ToM. More specifically, they identified seven categories of mental and social understanding: emotions, desires, intentions, precepts, knowledge, beliefs, and mentalistic understanding of non-literal communication; and these categories themselves contained a total of 39 ToM sub-abilities. Beaudoin and colleagues also considered the psychometric properties (internal consistency, inter-rater and test-retest reliability) of the reviewed tasks and noted that the majority of measures have insufficient empirically-validated psychometry, including poor replicability. The authors note that the great variety of ToM measures results in poor comparability across studies and adversely affects the dependability of the results. For instance, successful performance on FB paradigms may differ as a function of seemingly uncritical features of the experiment, such as how characters are presented (e.g., VR, line drawings, cartoons, text-based, etc.) and minute variations in the language used to present stimuli and/or instruct participants. The lack of standardisation and the resulting variability in effects concurs with the findings of the meta-analysis presented in Chapter 2, where substantial heterogeneity – even with studies employing the same paradigm – was found when examining the classically reported effect of age. To address heterogeneity and in the interest of bringing about greater standardisation of ToM protocols, Beaudoin et al. (2020) proposed the ATOMS (Abilities in Theory of Mind Space; Figure 5.1) framework which visually and logically categorises mental state and social understanding abilities.

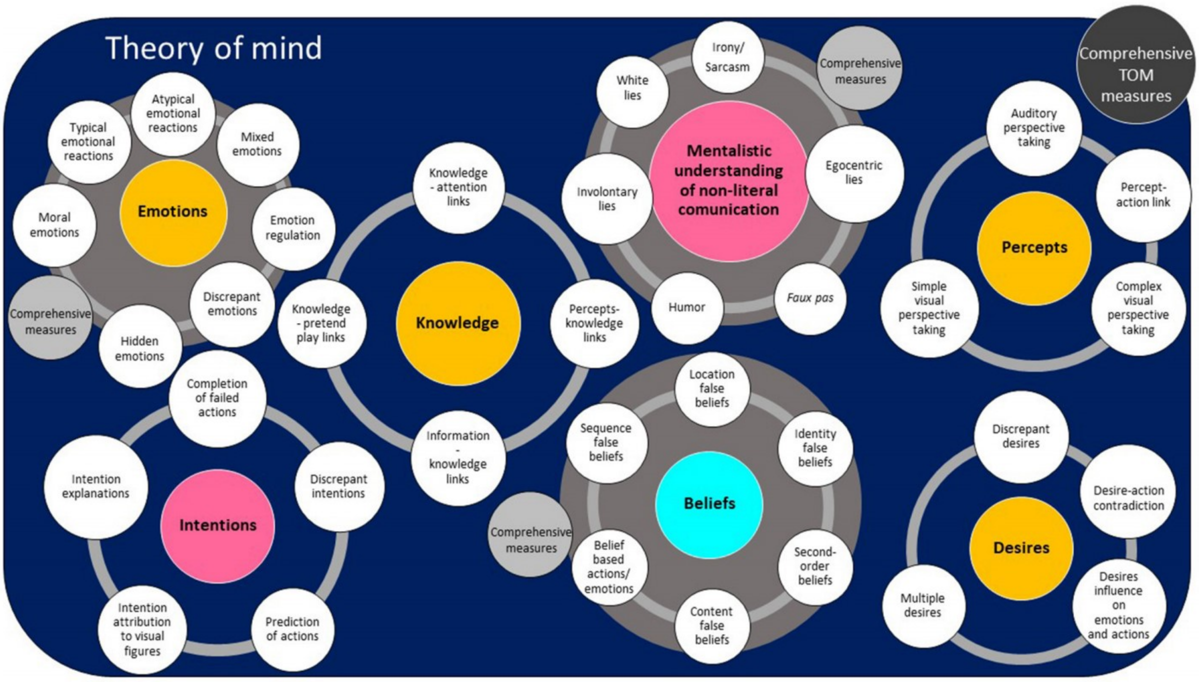


Figure 5.1. The ATOMS framework (Abilities in Theory of Mind Space): a visual representation of TOM categories and sub-abilities. Theory of mind space is represented as a large area that includes seven TOM categories of mental states and social situations understanding: Intentions, Desires, Emotions, Knowledge, Percepts, Beliefs, and mentalistic understanding of non-literal communication. Thirty-nine specific TOM sub-abilities (white circles) gravitate around the TOM category to which they pertain. Reproduced with permission from Beaudoin, Leblanc, Gagner, and Beauchamp (2020; *Front. Psychol. 10:2905*).

While a criticism of the framework is that some paradigms straddle multiple classifications and therefore do not fit neatly into the circumscribed "spaces", the framework is nevertheless a good starting point from which standardisation in terminology can be brought about. In the effort to bring about greater standardisation, recent work by Quesque and Rossetti (2020) is also important to consider. In their examination of the most commonly-used tests, Quesque and Rossetti reported that the majority of the reviewed ToM tasks do not require the representation of another's mental state and others fail to make the distinction between self and other processes. These two principles (mentalizing and self-other distinction or the non-merging principle), Quesque and Rossetti argue, are fundamental and only paradigms that satisfy these precepts should be classified as genuine tests of ToM. Our own novel FB tasks presented in Chapters 3 and 4 satisfy these two principles in that participants were required to infer (i.e., mentalize) about the agent's belief regarding the location of the target ball, and we manipulated compatibility between the participant's and agent's beliefs (i.e., SOK) which brought about a self-other distinction.

In addition to Beaudoin et al's ATOMS framework, Quesque and Rossetti's two criteria – termed mentalizing criterion and non-merging criterion – provide a foundation on which a clearer conceptualisation of ToM and its sub-abilities can be drawn. The framework and two criteria are helpful in providing both a terminological and theoretical (construct) clarification, and help define a limit or boundary as to what ToM actually means and as to what ToM tasks actually measure. It is our view, then, in integrating the work presented in this thesis with the recommendations by Beaudoin et al. (2020) and Quesque and Rossetti (2020), researchers should adhere to these principles both in the classification and implementation of ToM tasks.

Defining cognitive ageing

Normal cognitive ageing has been traditionally defined as changes that occur in individuals later in life that are free of overt disease and ill-health (Dumas, 2015). While a definition of cognitive ageing is relatively straightforward, what is less clear is when ageing *begins.* Many studies from the field of developmental psychology have shown that age-related changes occur across all stages of life, and that deterioration in some cognitive abilities can be detected in individuals in their mid-to-late twenties and prominent modifications in processing speed are observed from age 45 (Hoyer et al., 2004). Depending on the precise cognitive ability, age-related decay is evident far earlier than what is traditionally conceptualised as 'ageing'. In studies of ageing populations, there is great variation in what authors consider advanced age. This is an important consideration as the effect of age is greatly dependent on the average age of the older adult cohort, as we have seen in the meta-analysis presented in Chapter 2. Studies that employ *older* older adult samples (e.g., aged ≥ 80) may arbitrarily benefit from greater effect sizes (due to larger between-group differences) relative to studies that employ *younger* (e.g., aged 50-60) older adult samples. While both of these fictitious examples would be categorised as studies of ageing, a direct, uncorrected comparison of their results could lead to misleading interpretations. As we have seen in the meta-analysis in Chapter 2, studies of ageing ToM have employed a wide range of ages in their OA samples, and we have shown that variability in the age of the OA sample can modify the overall effect of age. What is more, some studies do not to distinguish between young-old and old-old; for example, an individual aged 61 may be classified in the same banding as another individual aged 81. Such studies are perhaps failing to detect nuanced and subtle alterations in cognition that occur across the entire lifespan, including *within* ageing. For these reasons, we suggest a definition of ageing as individuals aged 60 and above. Further, we suggest that studies make distinctions between young-old and old-old using the following convention: Young-Old for individuals aged 60-69, and Old-Old for individuals aged 70+. Through these conventions it is hoped that greater standardisation is achieved which may lead to more robust and replicable conclusions in the context of studies of healthy ageing.

However, it should be noted that the recommendations above should not be viewed as a fixed or static principle, and that these definitions may evolve over time. For example, there is currently a rapidly increasing percentage of senior citizens in the global population, and predictions suggest that the proportion of older adults relative to the wider population is set to increase (see Figure 5.2). Additionally, life expectancy has been steadily increasing globally for a number of decades, meaning that a greater proportion of the population will reach old (or, very old) age. In light of these trends, it is reasonable to suggest that our current conceptualisation of when ageing begins may need to shift to better reflect the underlying demographic changes in the global population.

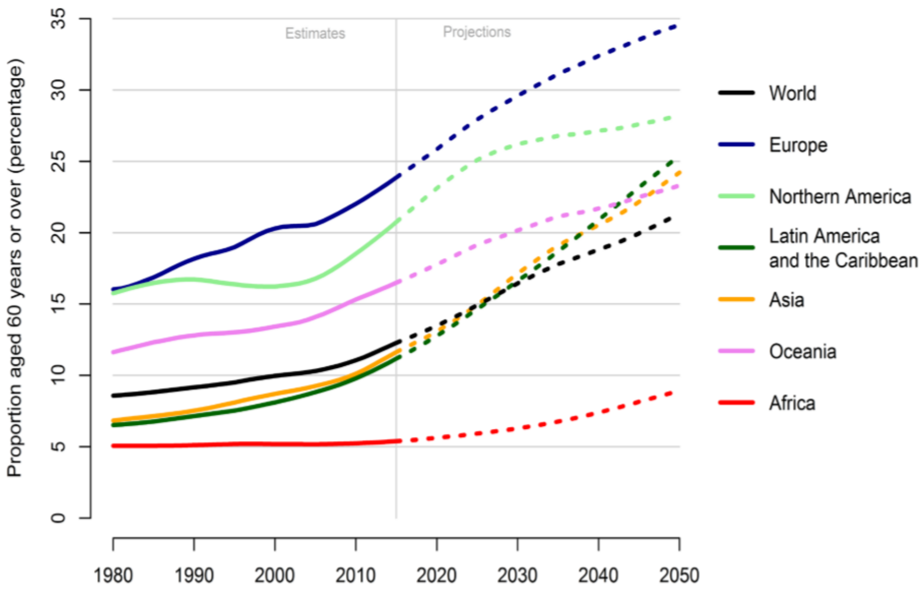


Figure 5.2. Percentage of the population aged 60 years or over by region, from 1980 to 2050. From "World Population Aging. Highlights" by Department of Economic and Social Affairs. United Nations (2017).

Non-social interference

We have shown that reasoning about FBs is difficult for OA beyond the non-social cognitive burdens of typical ToM studies. However, relative to their younger counterparts, we found that OA were particularly hindered by invalid cueing – a common feature of FB tasks. We also found that spatial working memory, a component of EF which in itself has been shown to decay with age (e.g., see Kumar & Priyadarshi, 2013), partially explained the cost associated with invalid (vs. valid) cueing. This is an important methodological consideration for past and present FB paradigms. Tasks that present multiple positions and agents require the processing of several different locations and their contents simultaneously and thus performance on these tasks may show age effects not exclusively because of impaired social cognitive function but perhaps also because of weakening EFs. A further but related issue is that the *degree* of difficulty – resulting from the above-mentioned non-social demands – is not uniform across tasks. This in part is driven by the somewhat amorphous nature of the concept of theory of mind, where no absolute or conclusive definition exists in the literature (see Quesque and Rossetti, 2020). Accordingly, as mentioned previously, traditionally, there have been no agreed upon boundaries or limits as to what the concept encompasses, and therefore there is substantial heterogeneity in paradigms and experimental protocols. Inconsistencies in task difficulty are found even within the *same* task type. Contrast Wimmer and Perner's (1983) classic object-location paradigm with Bernstein, Thornton, and Sommerville's (2011) sandbox task, where participants are asked to manually point to a location within a physical sandpit where a protagonist – holding a now invalid belief – is likely to search upon her return. The physical distance between the 'correct' location and the participant's indicated position, measured in centimetres, is the index by which ToM performance is determined. Superficially, both the classic object-location and the sandbox tasks examine FB understanding. However, it is doubtful that non-social processing demands are identical between these tasks, given that there are greater visuo-spatial and motor demands in visually orienting and estimating a psychical – as is the case in the Sandbox paradigm. Taken together, it is therefore plausible that the classically reported age effect in ToM has been somewhat inflated due to methodological confounds. i.e., additional processing demands that are supplementary to reasoning about others' minds. Therefore, this work has potentially far-reaching consequences as it raises questions as to how much of the classically reported age effect is indeed attributable to age-related changes in social cognitive function and how much is down to general age-related cognitive slowing.

Two-systems account of theory of mind

A key consideration in the interpretation of the current data is the two-system account of ToM (see Apperly & Butterfill, 2009). The two-system hypothesis posits that the computational nature of ToM processing can be either implicit (early-developing, automatic, rapid, and effortless) or explicit (late-developing, deliberate, slow, and effortful; Apperly & Butterfill, 2009; Butterfill & Apperly, 2013). The two systems are not mutually exclusive; indeed, many everyday activities require the simultaneous computation of both overt and covert mental states. However, the extent to which each of these two systems is employed in childhood, adulthood, and old age, and how the systems evolve over time, in addition to how they interact with individual differences in executive function and its decline in healthy ageing, is not yet fully understood (Schneider, Slaughter, & Dux, 2017). Indeed, the very presence of an *implicit* system is contested, as Kulke et al. (2018) recently showed in a systematic replication study where the authors concluded:

> [T]here is a strong need for systematic, large-scale, collaborative, and preregistered multilab replication and validation studies to explore more systematically whether implicit theory of mind is a real and robust phenomenon and under which conditions and in which age groups it can be measured. (Kulke et al., 2018, p. 11)

The experimental tasks employed in the two empirical investigations of this thesis rely on explicit and effortful mental state processing. The beliefs of the agents in the task are openly declared and the participant is required to hold in mind and reason about an overt belief state, and, therefore, according to the two-system theory, the current tasks require an *explicit* form of mentalizing. At the time of writing, no work currently exists comparing implicit vs explicit ToM reasoning in ageing, and so it is difficult to theorise on what effect—if any—ageing may have on implicit forms of mental state reasoning. In younger adults, research has shown that participants track the mental states of others even when they are not instructed to do so, and this tracking tends to rely less on EFs compared to explicit or effortful belief reasoning (Schneider, Nott, & Dux, 2014). Nevertheless, more work is needed to elucidate the relationship between implicit and explicit ToM in healthy ageing.

Limitations

There are a number of limitations to highlight in this study. Both empirical investigations presented in this thesis drew OA from an institutional database of participants aged

60+. The database, as is common with university participant pools, comprises largely a homogeneous sample. For example, in our study, OA disproportionately self-identified as ethnically white and over 50% were educated to at least under-graduate degree level – a rate that is higher than the general OA population in the United Kingdom. Due to the protective nature of education on certain cognitive functions, as we have seen in the meta-analysis in Chapter 2, variability in the levels of education in OA may potentially modify the magnitude age-related differences in ToM. For example, it is possible that greater age differences might have appeared with less well-educated OA samples. Still, we reported age differences in our behavioural data even with a highly-educated cohort of OAs and the level of education is comparable to other studies of ageing ToM (e.g., see Cho & Cohen, 2019). Nonetheless, greater efforts should be made to include more diverse and representative samples in ageing research.

Next, we originally planned to recruit and test 90 participants (45 young and 45 older) in total for the fMRI experiment reported in Chapter 4. The sample size calculation was based on a power analysis using Moran's (2013) ageing ToM data, where we aimed to achieve 90% statistical power in our design. However, as a result of the COVID-19 outbreak and its subsequent social restrictions, our recruitment was somewhat curtailed and only 72 participants were recruited: 50 YA and 22 OA. Of the 22 OA, only 19 had complete and usable data. A matched sub-set from the YA data was used in the final analysis, to allow for between-groups analyses. The final sample therefore comprised 38 participants, 19 young and 19 older adults. While we were able to replicate some of our initial behavioural findings and provide meaningful insights of the neural underpinnings of FB reasoning, our fMRI findings should be interpreted with caution due to low statistical power. It is of course possible that some of the reported null results may achieve statistical significance with a reasonably powered sample. An a posteriori power analysis to determine the actual achieved power was not performed as such post-hoc calculations have been shown to be logically invalid and practically misleading (Dziak, Dierker, & Abar, 2020).

Functional imaging, which was the basis of Chapter 4, has a number of methodological caveats. First, there are the classic limitations often associated with neuroimaging work, including the assumption of participant homogeneity and the subsequent processes involved cross-subject averaging. Also, as is traditional within the neuroimaging literature, we only recruited right-handed participants to combat potential effects arising from laterality. Further, due to the expansive exclusion criteria associated with MRI (e.g., cardiac implantable electronic devices, metallic foreign bodies, drug diffusion pumps, cochlear implants, etc.), only individuals who are generally fit and healthy are eligible to take part in f/MRI studies. As disease and general ill-health can potentially limit one's social experience, and by extension one's social and cognitive function, individuals who had poor general health (cardiovascular complaints, for example) were not included as to avoid confounds or systematic biases in the data. While one cannot say for certain that these results are generalisable to groups not included in this study, the sampling make-up is nonetheless reflective of the wider population, most of whom are 'neurotypical' and right-handed. On a more theoretical level, there are also limitations regarding causality and reverse inference with respect to fMRI, which is the tendency to theorise backwards from patterns of activity to deduce specific mental processes (e.g., see Poldrack, 2006). Other methods, such as virtual lesion approaches (e.g., transcranial magnetic stimulation [TMS]), offer the explanatory power of causality and allow researchers to test the direct, unbroken consequence of neuromodulation (*cause*) on mental processes (*effect*). On this basis, future investigations might employ neuromodulation approaches to investigate causally the role of the TPJ, for example, in supporting different features of FB reasoning (e.g., inhibiting the self-perspective), between young and older participants (e.g., see Hartwright et al., 2016 for an example of TMS modulation in the context of FB reasoning).

There is evidence to suggest that certain mental health conditions can modify ToM proficiency, including schizophrenia (see meta-analysis by Sprong et al., 2017) and depression (for recent example, see Washburn et al., 2016). Depression is of particular note as it has been shown to impair multiple cognitive domains, including EF, attention, and memory, (Rock, Roiser, Riedel, and Blackwell, 2014), all of which are important for a functioning ToM. Further, depression has been found to exacerbate executive dysfunction particularly in old age (Nakano et al., 2008; Pantzar et al., 2014). The current work in this thesis did not account for depression in terms of statistical adjustment and/or explicit screening—though, participants did self-declare no current diagnoses of neuropsychiatric disease. Given the evidence linking depression to cognitive dysfunction, particularly late life depression, it is plausible that depression was a contributing factor to the heterogeneity in ToM deficits seen in Chapter 2 in the meta-analysis. Future work might benefit from including measures of depression (or mental health disorders more generally) to address its potential contribution to ToM functioning, particularly in older adults.

Future directions

As is typical with scientific investigation, the research presented in this thesis has spawned many questions, perhaps more than it has answered. It is of course not feasible to outline *every* possible extension of the current work, so I will focus on the areas that are perhaps most pressing.

Developmental studies of age-related changes in socio-cognitive function have found detectable alterations in ToM abilities as early as the third decade of life (for example, see Figure 5.3 below). While changes are most pronounced in advanced age, studies employing ‘lifespan’ approaches have shown a linear downward trajectory from early-to-middle adulthood. Likewise, wider domain-general abilities, such as processing speed or memory, have also been shown to follow a similar developmental trajectory, see Figure 5.4.

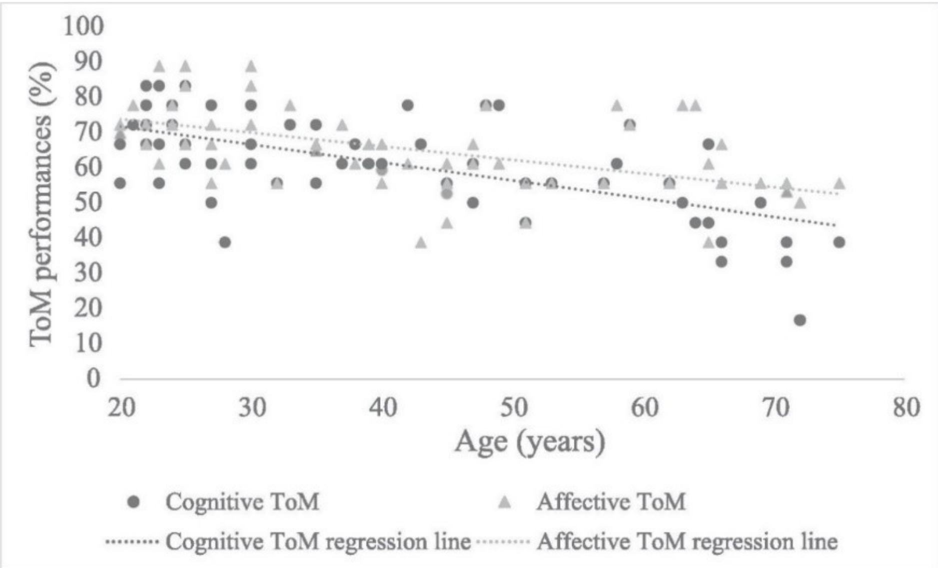


Figure 5.3. Cognitive and affective theory of mind performances across the adult lifespan. From Laillier et al. (2019) – reproduced with permission.

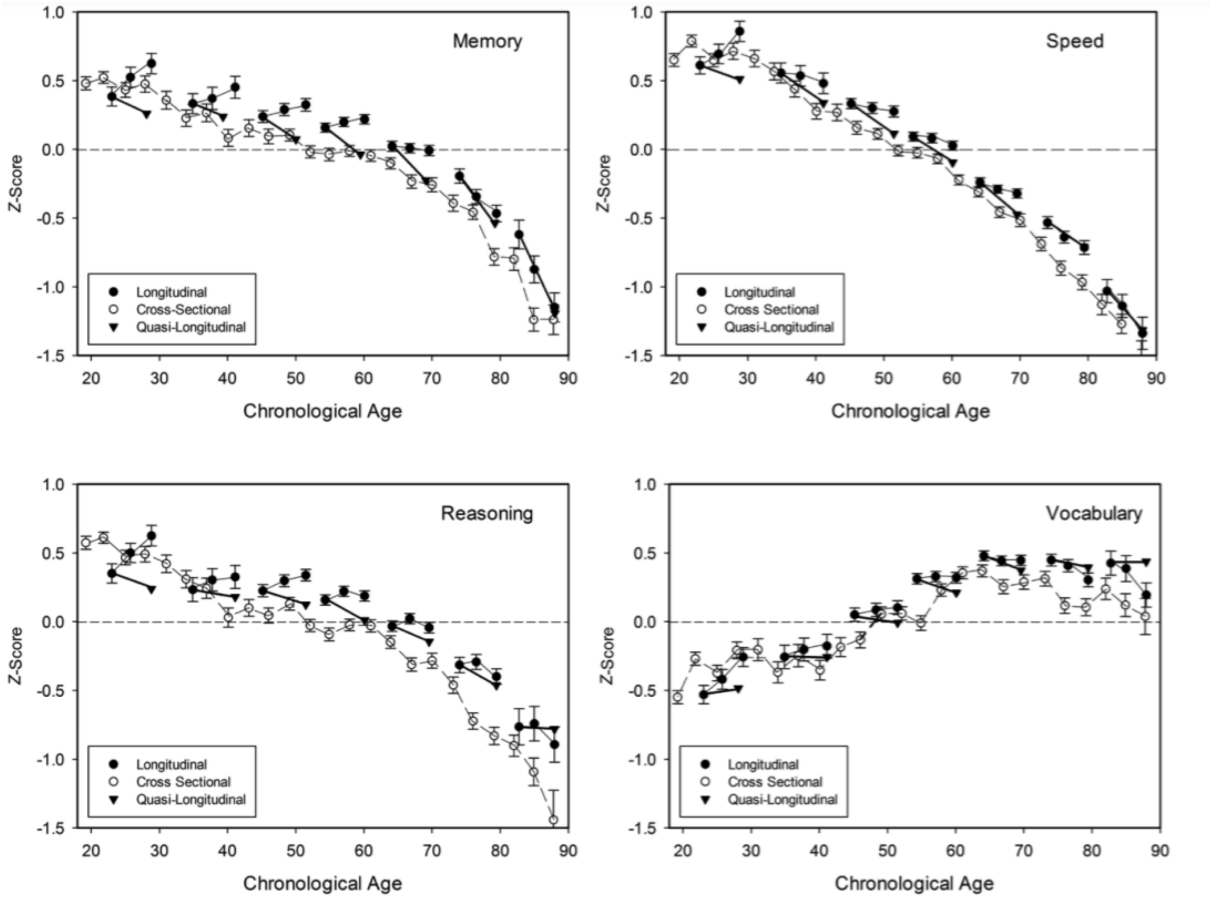


Figure 5.4. Trajectories of memory, processing speed, reasoning, and vocabulary against age. Note that vocabulary has been classically reported to *increase* (or at least, remain preserved) with age. From Salthouse (2019) – reproduced with permission.

It is important, therefore, to study a lifetime arc of social cognitive function which may provide greater insights into the precursors of group differences and ToM in advanced age. While there are examples in the extant literature of lifespan studies (e.g., Giovagnoli, 2019; Laillier et al., 2019) these are typically marred with issues surrounding low samples sizes and power. For example, Laillier et al. (2019) administered the MASC film paradigm to neurologically-healthy participants aged 20-75, with a mean age of 42. In total, 60 participants were recruited; while the authors do not report the frequencies and/or distribution of participant ages, there are still nevertheless fewer than 10 participants, on average, per age decile. Laillier and colleagues employed a regression-based approach and thus did not perform between-groups analyses; still, such small sample sizes in the context of lifespan studies is relatively small. Therefore, future studies investigating belief representation in ageing should include a more comprehensive sample, both in terms of age range – i.e., through a lifespan approach – and participant numbers.

Finally, the current investigation only employed *functional* approaches and did not utilise *structural* techniques, such as fibre/neural tractography or other volumetric methods. There is a body of evidence to suggest that individual differences in white matter integrity may account for variability in ToM performance. For example, Charlton et al. (2009) tested 106 older adults on a subset of Happé's Strange Stories (Happé et al., 1998) after which participants' structural brain data were acquired through diffusion-tensor imaging. The authors reported that performance on the stories task significantly correlated with white matter integrity even after controlling for performance on non-ToM conditions. Further, connectivity-based approaches have also shown that age-related changes in SCF can be traced back to dysregulated connectivity, especially in regions commonly thought to be associated with ToM – i.e., the *mentalizing network*. Hughes et al. (2019) recently tested 75 adults (40 young, 35 older) on Saxe and Kanwisher's (2003) well-known localizer task in addition to acquiring functional resting-state data. Participants' brain activity during resting-state scans were used as an index for intrinsic or baseline connectivity. Hughes and colleagues found that OA displayed poorer behavioural performance and weaker intrinsic connectivity within the mentalizing network compared to YA. Intrinsic connectivity between the right TPJ and the right temporal pole mediated age differences in ToM. More precisely, OA had weaker intrinsic connectivity between the right TPJ and right temporal pole that explained their poorer mentalizing behavioural performance. Taken together, it is important therefore to build a more comprehensive account of age-related changes in SCF by including both structural and functional neuroimaging approaches. While the omission of structural measures is indeed a limitation, additional data, including two resting state sequences, are currently being collected for the fMRI study presented in Chapter 4. These additional data will also be analysed using connectivity-based approaches, which was not possible with the current, incomplete data. Once data collection is complete, the full dataset will be written up separately and submitted for publication.

Conclusion

Social-cognitive abilities, and in particular ToM, are important for everyday social interactions and interpersonal relationships. While the early literature on ageing ToM produced somewhat contradictory findings, contemporary work has suggested that there is a general deterioration of ToM in advanced age. To bring greater clarity on an effect of age, and delineate the precise sources of difficulty when reasoning about others' beliefs in older age, three studies were presented in this thesis: one meta-analysis and two empirical investigations. In our quantitative review, we found that ToM competence does indeed seem to diminish with age, but importantly, the magnitude of this decline is modified – and in some cases, abolished – depending on task and sample considerations. Our behavioural and neuroimaging investigations showed that FB reasoning is effortful for OA beyond the non-social cognitive demands of classic ToM investigations. Individual differences in inhibitory control and spatial working memory partially explained some ToM performance in both young and older adults. However, the magnitude of conflict experienced and the cognitive systems used to resolve this were condition specific: managing competing cognitive perspectives was supported by attentional abilities whereas invalid cueing appeared to draw on spatial working memory. In our fMRI work, we again differentiated between possessing (vs. not possessing) a privileged knowledge of reality and found this to improve performance, and highlighted distinct regions – namely, the right TPJ and fusiform and cingulate gyri – involved in this process. We were also able to show specificity of the knowledge of reality effect within the anterior TPJ, a region thought to be associated with wider attention mechanisms. We further observed an interaction between knowledge of reality and self-other congruence, again both on a behavioural and neural level, with classic mentalizing nodes emerging as significant clusters, including the precuneus, TPJ, and STS involved in resolving incongruence in self-other beliefs when participants did not possess a privileged of reality. However, we did not observe differences based on a core manipulation of self-other congruence ([in]compatibility in self-other beliefs), either in the behavioural or neuroimaging data. While we showed emerging group differences in certain cortical regions, as informed by a ToM localizer, we did not observe statistically relevant effects or contrasts between young and older subjects in the neuroimaging data. In light of the evidence presented in this thesis, while we do not negate the presence of age-related decline of ToM abilities, we propose that an effect of age has been overestimated due to non-social executive demands embedded within typical ToM tasks. We recommend that researchers studying ageing ToM carefully construct tasks in a way that minimizes (or at least controls for) cognitive load beyond what is ordinarily required to reason about the beliefs of others.

References

Abdelkarim, D., Zhao, Y., Turner, M. P., Sivakolundu, D. K., Lu, H., & Rypma, B. (2019). A neural-vascular complex of age-related changes in the human brain: Anatomy, physiology, and implications for neurocognitive aging. *Neuroscience & Biobehavioral Reviews, 107*, 927-944.

Abu-Akel, A., & Shamay-Tsoory, S. (2011). Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia, 49*(11), 2971-2984.

Aizenstein, H. J., Clark, K. A., Butters, M. A., Cochran, J., Stenger, V. A., Meltzer, C. C., ... & Carter, C. S. (2004). The BOLD hemodynamic response in healthy aging. *Journal of Cognitive Neuroscience, 16*(5), 786-793.

Alain, C., & Woods, D. L. (1999). Age-related changes in processing auditory stimuli during visual attention: evidence for deficits in inhibitory control and sensory memory. *Psychology and Aging, 14*(3), 507.

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.

Aminoff, E. M., Kveraga, K., & Bar, M. (2013). The role of the parahippocampal cortex in cognition. *Trends in Cognitive Sciences, 17*(8), 379-390.

Apperly, I. A. (2010). *Mindreaders: The Cognitive Basis of ‘Theory of Mind’*. Hove: Psychology Press.

Apperly, I. A. (2012). What is “theory of mind”? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology, 65*(5), 825-839.

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. doi: 10.1016/j.cognition.2007.05.005.

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*(5), 1691-1703. doi:10.1111/j.1467-8624.2011.01635.x.

Ashinoff, B. K., Geng, J. J., & Mevorach, C. (2019). Delayed reactive distractor suppression in aging populations. *Psychology and Aging, 34*(3), 418.

Astington, J. W. (1993). *The child's discovery of the mind* (Vol. 31). Harvard University Press.

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective. *Frontiers in Psychology, 5*, 655. doi:10.3389/fpsyg.2014.00655.

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., & Von Hippel, W. (2008). Empathy and social functioning in late adulthood. *Aging and Mental Health, 12*(4), 499-503.

Baksh, R. A., Bugeja, T., & MacPherson, S. E. (2020). Executive functions do not underlie performance on the Edinburgh Social Cognition Test (ESCoT) in healthy younger and older adults. *Journal of the International Neuropsychological Society*, *26*(5), 527-538.

Baksh, R. A., Abrahams, S., Auyeung, B., & MacPherson, S. E. (2018). The Edinburgh Social Cognition Test (ESCoT): Examining the effects of age on a new measure of theory of mind and social norm understanding. *PloS One, 13*(4), e0195818.

Banich, M. T., & Depue, B. E. (2015). Recent advances in understanding neural systems that support inhibitory control. *Current Opinion in Behavioral Sciences, 1*, 17-22.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. *Cognition, 21*(1), 37-46.

Beadle, J. N., Paradiso, S., Kovach, C., Polgreen, L., Denburg, N., & Tranel, D. (2012). Effects of age-related differences in empathy on social economic decision-making. *International Psychogeriatrics/IPA*, 24(5), 822.

Begg, C. B. (1994). Publication bias. The handbook of research synthesis, 25, 299-409.

Bendayan, R., Piccinin, A. M., Hofer, S. M., Cadar, D., Johansson, B., & Muniz-Terrera, G. (2017). Decline in memory, visuospatial ability, and crystalized cognitive abilities in older adults: normative aging or terminal decline?. *Journal of Aging Research*, *2017*.

Bernstein, D. M., Erdfelder, E., Meltzoff, A. N., Peria, W., & Loftus, G. R. (2011). Hindsight bias from 3 to 95 years of age. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 378.

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

Birch, S. A., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences, 8*(6), 255-260. doi:10.1016/j.tics.2004.04.011.

Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*(5), 382-386. doi:10.1111/j.1467-9280.2007.01909.x.

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 60*(5), P223-P233.

Bora, E., & Berk, M. (2016). Theory of mind in major depressive disorder: A meta-analysis. *Journal of Affective Disorders, 191*, 49-55.

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.

Brown, L. A., Brockmole, J. R., Gow, A. J., & Deary, I. J. (2012). Processing speed and visuospatial executive function predict visual working memory ability in older adults. *Experimental Aging Research, 38*(1), 1-19.

Brunet, E., Sarfati, Y., Hardy-Baylé, M. C., & Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal task. *Neuroimage, 11*(2), 157-166.

Bugg, J. M., Zook, N. A., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2006). Age differences in fluid intelligence: Contributions of general slowing and frontal decline. *Brain and Cognition, 62*(1), 9-16.

Caballero, H. S., McFall, G. P., Wiebe, S. A., & Dixon, R. A. (2020). Integrating Three Characteristics of Executive Function in Non-Demented Aging: Trajectories, Classification, and Biomarker Predictors. *Journal of the International Neuropsychological Society*, 1-14.

Cabeza, R., Albert, M., Belleville, S., Craik, F. I., Duarte, A., Grady, C. L., ... & Rajah, M. N. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience, 19*(11), 701-710.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage, 17*(3), 1394-1402.

Cabeza, R., & Dennis, N. A. (2012). Frontal lobes and aging. *Principles of Frontal Lobes Function, 628-652.*

Calso, C., Besnard, J., & Allain, P. (2019). Frontal lobe functions in normal aging: metacognition, autonomy, and quality of life. *Experimental Aging Research, 45*(1), 10-27.

Calso, C., Besnard, J., & Allain, P. (2020). Study of the theory of mind in normal aging: focus on the deception detection and its links with other cognitive functions. *Aging, Neuropsychology, and Cognition, 27*(3), 430-452.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development: An International Journal of Research and Practice, 11*(2), 73-92.

Cattell, R. B., & Horn, J. L. (1978). A check on the theory of fluid and crystallised intelligence with description of new subtest designs. *Journal of Educational Measurement, 15*(3), 139-164.

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. *The International Journal of Aging and Human Development, 76*(3), 181-198.

Chalah, M. A., Kauv, P., Lefaucheur, J. P., Hodel, J., Créange, A., & Ayache, S. S. (2017). Theory of mind in multiple sclerosis: a neuropsychological and MRI study. *Neuroscience Letters, 658*, 108-113.

Choi, H. J., Lee, D. Y., Seo, E. H., Jo, M. K., Sohn, B. K., Choe, Y. M., ... & Woo, J. I. (2014). A normative study of the digit span in an educationally diverse elderly population. *Psychiatry Investigation, 11*(1), 39.

Craik, F. I., & Salthouse, T. A. (Eds.). (2011). *The handbook of aging and cognition*. Psychology Press.

Davis, M.H. (1980). A multidimensional approach to individual differences in empathy. *Catalog of Selected Documents in Psychology, 10*(4), 1-17.

DeDe, G., & Flax, J. K. (2016). Language comprehension in aging. *Cognition, Language and Aging*, 107-133.

Dennett, D. C. (1978). Toward a cognitive theory of consciousness.

Desgranges, B., Laisney, M., Bon, L., Duval, C., Mondou, A., Bejanin, A., ... & Muckle, G. (2012). TOM-15: Une épreuve de fausses croyances pour évaluer la théorie de l'esprit cognitive. *Revue de Neuropsychologie, 4*(3), 216-220.

Duval, C., Piolino, P., Bejanin, A., Eustache, F., & Desgranges, B. (2011). Age effects on different components of theory of mind. *Consciousness and cognition, 20*(3), 627-642.

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., ... & Convit, A. (2006). Introducing MASC: a movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders, 36*(5), 623-636.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629-634.

El Haj, M., Raffard, S., & Gély-Nargeot, M. C. (2016). Destination memory and cognitive theory of mind in normal ageing. *Memory, 24*(4), 526-534.

Eyler, L. T., Sherzai, A., Kaup, A. R., & Jeste, D. V. (2011). A review of functional brain imaging correlates of successful cognitive aging. *Biological Psychiatry, 70*(2), 115-122.

Ferguson, H. J., Apperly, I., & Cane, J. E. (2017). Eye tracking reveals the cost of switching between self and other perspectives in a visual perspective-taking task. *Quarterly Journal of Experimental Psychology, 70*(8), 1646-1660.

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

Ferguson, H. J., Brunsdon, V. E., & Bradford, E. E. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports, 11*(1), 1-17.

Fink, E., Begeer, S., Peterson, C. C., Slaughter, V., & de Rosnay, M. (2015). Friendlessness and theory of mind: A prospective longitudinal study. *British Journal of Developmental Psychology, 33*(1), 1-17.

Fischer, A. L., O'Rourke, N., & Loken Thornton, W. (2017). Age differences in cognitive and affective theory of mind: Concurrent contributions of neurocognitive performance, sex, and pulse pressure. *The Journals of Gerontology: Series B, 72*(1), 71-81.

Fjell, A. M., Sneve, M. H., Grydeland, H., Storsve, A. B., & Walhovd, K. B. (2017). The disconnected brain and executive function decline in aging. *Cerebral Cortex, 27*(3), 2303-2317.

Flynn, E., O'Malley, C., & Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science, 7*(1), 103-115.

Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E. J., Derrfuss, J., Imperati, D., & Brown, S. (2011). The speed-accuracy tradeoff in the elderly brain: a structural model-based approach. *Journal of Neuroscience, 31*(47), 17242-17249.

Franco, M. G., & Smith, P. K. (2013). Theory of mind, old age, and educational opportunity in Colombia. *The International Journal of Aging and Human Development, 77*(2), 107-125.

Freedman, A., & Nicolle, J. (2020). Social isolation and loneliness: The new geriatric giants: Approach for primary care. *Canadian Family Physician, 66*(3), 176-182.

Friedman, O., & Leslie, A. M. (2004). Mechanisms of belief-desire reasoning: Inhibition and bias. *Psychological Science, 15*(8), 547-552.

Friedman, O., & Leslie, A. M. (2005). Processing demands in belief - desire reasoning: Inhibition or general difficulty?. *Developmental Science*, *8*(3), 218-225. doi: 10.1111/j.1467-7687.2005.00410.

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology, 15*(17), R644-R645.

Gazzaley, A., Cooney, J. W., Rissman, J., & D'esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience, 8*(10), 1298-1300.

Geng, J. J., & Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: contextual updating?. *Neuroscience & Biobehavioral Reviews, 37*(10), 2608-2620.

German, T., & Hehman, J. (2006). Representational and executive selection resources in 'theory of mind': Evidence from compromised belief-desire reasoning in old age. *Cognition, 101*(1), 129-152. doi:10.1016/j.cognition.2005.05.007.

Ghrear, S. E., Birch, S. A., & Bernstein, D. M. (2016). Outcome knowledge and false belief. *Frontiers in Psychology, 7*, 118. doi:10.3389/fpsyg.2016.00118.

Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience, 13*(7), 491-505.

Grady, C. L., Maisog, J. M., Horwitz, B., Ungerleider, L. G., Mentis, M. J., Salerno, J. A., ... & Haxby, J. V. (1994). Age-related changes in cortical blood flow activation during visual processing of faces and location. *Journal of Neuroscience, 14*(3), 1450-1462.

Grainger, S. A., Henry, J. D., Naughtin, C. K., Comino, M. S., & Dux, P. E. (2018). Implicit false belief tracking is preserved in late adulthood. *Quarterly Journal of Experimental Psychology, 71*(9), 1980-1987.

Grainger, S. A., Rakunathan, V., Adams, A. G., Canty, A. L., & Henry, J. D. (2020). An assessment of age differences in theory of mind using the virtual assessment of mentalizing ability. *Aging, Neuropsychology, and Cognition, 28*(1), 97-107.

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.

Haidich, A. B. (2010). Meta-analysis in medical research. *Hippokratia, 14*(Suppl 1), 29.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24*(2), 129-154.

Happé, F. G., Winner, E., & Brownell, H. (1998). The getting of wisdom: theory of mind in old age. *Developmental Psychology, 34*(2), 358.

Harada, C. N., Love, M. C. N., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine*, *29*(4), 737-752.

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. doi:10.1016/j.neuroimage.2012.03.012.

Hartwright, C. E., Apperly, I. A., & Hansen, P. C. (2013). Representation, Control, or Reasoning? Distinct Functions for Theory of Mind within the Medial Prefrontal Cortex. *Journal of Cognitive Neuroscience, 26*(4), 683-698. doi:10.1162/jocn_a_00520.

Hartwright, C. E., Apperly, I. A., & Hansen, P. C. (2015). The special case of self-perspective inhibition in mental, but not non-mental, representation. *Neuropsychologia, 67*, 183-192. doi:10.1016/j.neuropsychologia.2014.12.015.

Hartwright, C. E., Hardwick, R. M., Apperly, I. A., & Hansen, P. C. (2016). Resting state morphology predicts the effect of theta burst stimulation in false belief reasoning. *Human Brain Mapping*, *37*(10), 3502-3514. doi: 10.1002/hbm.23255.

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. doi: 10.1037/a0030677.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557-560.

Hoaglin, D. C. (2016). Misunderstandings about Q and 'Cochran's Q test'in meta - analysis. *Statistics in Medicine, 35*(4), 485-495.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallised intelligence. *Acta Psychologica, 26*, 107-129.

Hsieh, S., & Lin, Y. C. (2017). Strategies for stimulus selective stopping in the elderly. *Acta Psychologica*, *173*, 122-131.

Hsieh, S., McGrory, S., Leslie, F., Dawson, K., Ahmed, S., Butler, C. R., ... & Hodges, J. R. (2015). The Mini-Addenbrooke's Cognitive Examination: A new assessment tool for dementia. *Dementia and Geriatric Cognitive Disorders*, *39*(1-2), 1-11. doi: 10.1159/000366040.

Huettel, S. A., Singerman, J. D., & McCarthy, G. (2001). The effects of aging upon the hemodynamic response measured by functional MRI. *Neuroimage, 13*(1), 161-175.

Hughes, C., Cassidy, B. S., Faskowitz, J., Avena-Koenigsberger, A., Sporns, O., & Krendl, A. C. (2019). Age differences in specific neural connections within the Default Mode Network underlie theory of mind. *NeuroImage, 191*, 269-277.

Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, *43*(6), 1447.

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta - analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment, 8*(4), 275-292.

Ibanez, A., Huepe, D., Gempp, R., Gutiérrez, V., Rivera-Rei, A., & Toledo, M. I. (2013). Empathy, sex and fluid intelligence as predictors of theory of mind. *Personality and Individual Differences, 54*(5), 616-621.

Ikram, M. A., Vrooman, H. A., Vernooij, M. W., van der Lijn, F., Hofman, A., van der Lugt, A., ... & Breteler, M. M. (2008). Brain tissue volumes in the general elderly population: The Rotterdam Scan Study. *Neurobiology of Aging, 29*(6), 882-890.

Ioannidis, J. P. (2017). Meta-analyses can be credible and useful: a new standard. *JAMA Psychiatry, 74*(4), 311-312.

Isingrini, M., & Vazou, F. (1997). Relation between fluid intelligence and frontal lobe functioning in older adults. *The International Journal of Aging and Human Development, 45*(2), 99-109.

Johansson Nolaker, E., Murray, K., Happé, F., & Charlton, R. A. (2018). Cognitive and affective associations with an ecologically valid test of theory of mind across the lifespan*. Neuropsychology, 32*(6), 754.

Kalbe, E., Schlegel, M., Sack, A. T., Nowak, D. A., Dafotakis, M., Bangard, C., ... & Kessler, J. (2010). Dissociating cognitive from affective theory of mind: a TMS study. *Cortex, 46*(6), 769-780.

Kanske, P., Böckler, A., Trautwein, F. M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain–behavior relations for empathy and Theory of Mind. *NeuroImage, 122*, 6-19.

Kaufman, A. S., Kaufman, J. C., Liu, X., & Johnson, C. K. (2009). How do educational attainment and gender relate to fluid intelligence, crystallised intelligence, and academic skills at ages 22–90 years?. *Archives of Clinical Neuropsychology, 24*(2), 153-163.

Kaufman, A. S., & Lichtenberger, E. O. (2005). Assessing adolescent and adult intelligence. John Wiley & Sons.

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging, 16*(2), 312.

Kievit, R. A., Davis, S. W., Mitchell, D. J., Taylor, J. R., Duncan, J., & Henson, R. N. (2014). Distinct aspects of frontal lobe structure mediate age-related differences in fluid intelligence and multitasking. *Nature Communications, 5*(1), 1-10.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1-2), 262-284.

Kronbichler, L., Tschernegg, M., Martin, A. I., Schurz, M., & Kronbichler, M. (2017). Abnormal brain activation during theory of mind tasks in schizophrenia: a meta-analysis. *Schizophrenia Bulletin, 43*(6), 1240-1250.

Kulke, L., von Duhn, B., Schneider, D., & Rakoczy, H. (2018). Is implicit theory of mind a real and robust phenomenon? Results from a systematic replication study. *Psychological Science, 29*(6), 888-900.

Laillier, R., Viard, A., Caillaud, M., Duclos, H., Bejanin, A., de La Sayette, V., ... & Laisney, M. (2019). Neurocognitive determinants of theory of mind across the adult lifespan. *Brain and Cognition, 136*, 103588.

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.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences, 8*(12), 528-533. doi:10.1016/j.tics.2004.10.001.

Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology, 50*(1), 45-85. doi:10.1016/j.cogpsych.2004.06.002.

Li, R., Yin, S., Zhu, X., Ren, W., Yu, J., Wang, P., ... & Li, J. (2017). Linking inter-individual variability in functional brain connectivity to cognitive ability in elderly individuals. *Frontiers in Aging Neuroscience, 9*, 385.

Li, X., Wang, K., Wang, F., Tao, Q., Xie, Y., & Cheng, Q. (2013). Aging of theory of mind: The influence of educational level and cognitive processing. *International Journal of Psychology, 48*(4), 715-727.

Liu, R., Patel, B. N., & Kwon, M. (2017). Age-related changes in crowding and reading speed. *Scientific Reports, 7*(1), 1-10.

Love, M.C.N. (2015). Social cognition in older adults: a review of neuropsychology, neurobiology, and functional connectivity. *ARCHIVOS DE MEDICINA 1*, 6.

Madden, D. J., Gottlob, L. R., Denny, L. L., Turkington, T. G., Provenzale, J. M., Hawk, T. C., & Coleman, R. E. (1999). Aging and recognition memory: Changes in regional cerebral blood flow associated with components of reaction time distributions. *Journal of Cognitive Neuroscience, 11*(5), 511-520.

Maldonado, T., Orr, J. M., Goen, J. R., & Bernard, J. A. (2020). Age differences in the subcomponents of executive functioning. *The Journals of Gerontology: Series B, 75*(6), e31-e55.

Manard, M., Bahri, M. A., Salmon, E., & Collette, F. (2016). Relationship between grey matter integrity and executive abilities in aging. *Brain Research, 1642*, 562-580.

Marcovitch, S., O'Brien, M., Calkins, S. D., Leerkes, E. M., Weaver, J. M., & Levine, D. W. (2015). A longitudinal assessment of the relation between executive function and theory of mind at 3, 4, and 5 years. *Cognitive Development, 33*, 40-55.

Mars, R. B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., & Rushworth, M. F. (2012). Connectivity-based subdivisions of the human right "temporoparietal junction area": evidence for different areas participating in different cortical networks. *Cerebral Cortex, 22*(8), 1894-1903. doi: 10.1093/cercor/bhr268.

Martín-Rodríguez, J. F., & León-Carrión, J. (2010). Theory of mind deficits in patients with acquired brain injury: A quantitative review. *Neuropsychologia, 48*(5), 1181-1191.

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.

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology, 38*(1), 115.

Mevorach, C., Spaniol, M. M., Soden, M., & Galea, J. M. (2016). Age-dependent distractor suppression across the vision and motor domain. *Journal of Vision*, *16*(11), 27-27.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167-202.

Mitchell, P., Robinson, E. J., Isaacs, J. E., & Nye, R. M. (1996). Contamination in reasoning about false belief: An instance of realist bias in adults but not children. *Cognition, 59*(1), 1-21.

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.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine, 6*(7), e1000097.

Moran, J. M. (2013). Lifespan development: The effects of typical aging on theory of mind. *Behavioural Brain Research*, *237*, 32-40.

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

Morou, N., Papaliagkas, V., Markouli, E., Karagianni, M., Nazlidou, E., Spilioti, M., ... & Kosmidis, M. H. (2018). Theory of Mind impairment in focal versus generalized epilepsy. *Epilepsy & Behavior, 88*, 244-250.

Murray, K., Johnston, K., Cunnane, H., Kerr, C., Spain, D., Gillan, N., ... & Happé , F. (2017). A new test of advanced theory of mind: The "Strange Stories Film Task" captures social processing differences in adults with autism spectrum disorders. *Autism Research, 10*(6), 1120-1132.

Mutter, B., Alcorn, M. B., & Welsh, M. (2006). Theory of mind and executive function: Working-memory capacity and inhibitory control as predictors of false-belief task performance. *Perceptual and Motor Skills, 102*(3), 819-835.

Nolaker, E., Murray, K., Happé , F., & Charlton, R. A. (2018). Cognitive and affective associations with an ecologically valid test of theory of mind across the lifespan. *Neuropsychology, 32*(6), 754.

Opdebeeck, C., Martyr, A., & Clare, L. (2016). Cognitive reserve and cognitive function in healthy older people: a meta-analysis. *Aging, Neuropsychology, and Cognition, 23*(1), 40-60.

Pantell, M., Rehkopf, D., Jutte, D., Syme, S. L., Balmes, J., & Adler, N. (2013). Social isolation: a predictor of mortality comparable to traditional clinical risk factors. *American Journal of Public Health, 103*(11), 2056-2062.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annual Review of Psychology, 60*, 173-196.

Perner, J., & Lang, B. (2000). *Theory of mind and executive function: Is there a developmental relationship?* In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen

(Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (p. 150–181). Oxford University Press.

Perner, J., Lang, B., & Kloo, D. (2002). Theory of mind and self‐control: More than a common problem of inhibition. *Child Development, 73*(3), 752-767.

Peterson, C. C., & Siegal, M. (2002). Mindreading and moral awareness in popular and rejected preschoolers. *British Journal of Developmental Psychology, 20*(2), 205-224.

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. doi:https://doi.org/10.1016/j.cognition.2011.05.003.

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

Pylyshyn, Z. W. (1978). When is attribution of beliefs justified?[P&W]. *Behavioral and Brain Sciences, 1*(4), 592-593.

Quesque, F., & Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. *Perspectives on Psychological Science, 15*(2), 384-396.

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.

Rakoczy, H., Wandt, R., Thomas, S., Nowak, J., & Kunzmann, U. (2018). Theory of mind and wisdom: The development of different forms of perspective‐taking in late adulthood. *British Journal of Psychology, 109*(1), 6-24.

Raz, N., Gunning-Dixon, F. M., Head, D., Dupuis, J. H., & Acker, J. D. (1998). Neuroanatomical correlates of cognitive aging: evidence from structural magnetic resonance imaging. *Neuropsychology, 12*(1), 95.

Reiter, A. M., Kanske, P., Eppinger, B., & Li, S. C. (2017). The aging of the social mind-differential effects on components of social understanding. *Scientific Reports*, *7*(1), 1-8.

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. A., & Park, D. C. (2014). How does it STAC up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychology Review, 24*(3), 355-370.

Rey-Mermet, A., & Gade, M. (2018). Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin & Review, 25*(5), 1695-1716.

Rico-Uribe, L. A., Caballero, F. F., Olaya, B., Tobiasz-Adamczyk, B., Koskinen, S., Leonardi, M., ... & Miret, M. (2016). Loneliness, social networks, and health: a cross-sectional study in three countries. *PloS One, 11*(1), e0145264.

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science, 29*(8), 1358-1369.

Ronchi, L., Banerjee, R., & Lecce, S. (2020). Theory of mind and peer relationships: The role of social anxiety. *Social Development, 29*(2), 478-493.

Ruitenberg, M. F., Santens, P., & Notebaert, W. (2020). Cognitive and Affective Theory of Mind in Healthy Aging. *Experimental Aging Research, 46*(5), 382-395.

Sakyi, K. S., Surkan, P. J., Fombonne, E., Chollet, A., & Melchior, M. (2015). Childhood friendships and psychological difficulties in young adulthood: an 18-year follow-up study. *European Child & Adolescent Psychiatry, 24*(7), 815-826.

Salthouse, T. A. (1992). Influence of processing speed on adult age differences in working memory. *Acta Psychologica, 79*(2), 155-170.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403.

Salthouse, T. A., Fristoe, N., McGuthry, K. E., & Hambrick, D. Z. (1998). Relation of task switching to speed, age, and fluid intelligence. *Psychology and Aging, 13*(3), 445.

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science, 13*(4), 140-144.

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

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society, 16*(5), 754-760.

Salthouse, T. A. (2012). Consequences of age-related cognitive declines. *Annual Review of Psychology, 63*, 201-226.

Salthouse, T. A. (2019). Trajectories of normal cognitive aging*. Psychology and Aging, 34*(1), 17.

Saltzman, J., Strauss, E., Hunter, M., & Archibald, S. (2000). Theory of mind and executive functions in normal human aging and Parkinson's disease. *Journal of the International Neuropsychological Society, 6*(7), 781-788.

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: a case of a selective deficit in inhibiting self-perspective. *Brain, 128*(Pt 5), 1102-1111. doi:10.1093/brain/awh464.

Samson, D., Houthuys, S., & Humphreys, G. W. (2015). Self-perspective inhibition deficits cannot be explained by general executive control difficulties. *Cortex, 70*, 189-201. doi:10.1016/j.cortex.2014.12.021.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. *Neuroimage, 19*(4), 1835-1842. doi:10.1016/s1053-8119(03)00230-1.

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science*, *17*(8), 692-699. doi: 10.1016/j.tics.2018.04.006.

Schneider, B. A., Pichora-Fuller, K., & Daneman, M. (2010). Effects of senescent changes in audition and cognition on spoken language comprehension. In *The aging auditory system* (pp. 167-210). Springer, New York, NY.

Schneider, D., Slaughter, V. P., Bayliss, A. P., & Dux, P. E. (2013). A temporally sustained implicit theory of mind deficit in autism spectrum disorders. *Cognition, 129*(2), 410-417.

Schneider, D., Slaughter, V. P., & Dux, P. E. (2017). Current evidence for automatic Theory of Mind processing in adults. *Cognition, 162*, 27-31.

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E. N., & Saxe, R. (2009). Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. *PloS One, 4*(3), e4869. doi: 10.1371/journal.pone.0004869.

Schünemann, H., Brożek, J., Guyatt, G., & Oxman, A. (2013). *GRADE Handbook*. Available at: https://gdt.gradepro.org/app/handbook/handbook.html. [Accessed 16 Jan 2021].

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 9-34.

Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences, 21*(4), 237-249.

Sebastian, C. L., Fontaine, N. M., Bird, G., Blakemore, S. J., De Brito, S. A., McCrory, E. J., & Viding, E. (2012). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Social Cognitive and Affective Neuroscience, 7*(1), 53-63.

Segaert, K., Lucas, S. J. E., Burley, C. V., Segaert, P., Milner, A. E., Ryan, M., & Wheeldon, L. (2018). Higher physical fitness levels are associated with less language decline in healthy ageing. *Scientific Reports, 8*(1), 1-10.

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia, 45*(13), 3054-3067.

Shao, J., Zhang, Q., Ren, Y., Li, X., & Lin, T. (2019). Why are older adults victims of fraud? Current knowledge and prospects regarding older adults' vulnerability to fraud. *Journal of Elder Abuse & Neglect, 31*(3), 225-243.

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta - analysis of theory of mind and peer popularity in the preschool and early school years. *Child Development, 86*(4), 1159-1174.

Slessor, G., Phillips, L. H., & Bull, R. (2007). Exploring the specificity of age-related differences in theory of mind tasks. *Psychology and Aging, 22*(3), 639.

Smith, K. J., & Victor, C. (2019). Typologies of loneliness, living alone and social isolation, and their associations with physical and mental health. *Ageing & Society, 39*(8), 1709-1730.

Slessor, G., Phillips, L. H., & Bull, R. (2007). Exploring the specificity of age-related differences in theory of mind tasks. *Psychology and Aging, 22*(3), 639.

Sprong, M., Schothorst, P., Vos, E., Hox, J., & Van Engeland, H. (2007). Theory of mind in schizophrenia: meta-analysis. *The British Journal of Psychiatry, 191*(1), 5-13.

Staff, R. T., Hogan, M. J., & Whalley, L. J. (2014). Aging trajectories of fluid intelligence in late life: The influence of age, practice and childhood IQ on Raven's Progressive Matrices. *Intelligence, 47*, 194-201.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10*(5), 640-656.

Sullivan, S., & Ruffman, T. (2004). Social understanding: How does it fare with advancing years?. *British Journal of Psychology*, *95*(1), 1-18.

Tager-Flusberg, H. (2007). Evaluating the theory-of-mind hypothesis of autism. *Current Directions in Psychological Science, 16*(6), 311-315.

Terry, R. D., & Katzman, R. (2001). Life span and synapses: will there be a primary senile dementia?. *Neurobiology of Aging, 22*(3), 347-348.

Thornton, R., & Light, L. L. (2006). Language comprehension and production in normal aging. In *Handbook of the Psychology of Aging* (pp. 261-287). Academic Press.

Tsvetanov, K. A., Henson, R. N., & Rowe, J. B. (2021). Separating vascular and neuronal effects of age on fMRI BOLD signals. *Philosophical Transactions of the Royal Society B, 376*(1815), 20190631.

Tsvetanov, K. A., Henson, R. N., Tyler, L. K., Davis, S. W., Shafto, M. A., Taylor, J. R., ... & Rowe, J. B. (2015). The effect of ageing on f MRI: Correction for the confounding effects of vascular reactivity evaluated by joint f MRI and MEG in 335 adults. *Human Brain Mapping, 36*(6), 2248-2269.

Uddin, L. Q. (2013). Complex relationships between structural and functional brain connectivity. *Trends in Cognitive Sciences, 17*(12), 600-602.

Valenzuela, M. J., & Sachdev, P. (2006). Brain reserve and dementia: a systematic review. *Psychological Medicine, 36*(4), 441.

Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E., & Kliegel, M. (2013). Development of affective theory of mind across adolescence: disentangling the role of executive functions. *Developmental Neuropsychology, 38*(2), 114-125.

Washburn, D., Wilson, G., Roes, M., Rnic, K., & Harkness, K. L. (2016). Theory of mind in social anxiety disorder, depression, and comorbid conditions. *Journal of Anxiety Disorders, 37*, 71-77.

Wechsler, D. (1997). Wechsler memory scale (WMS-III) (Vol. 14). San Antonio, TX: Psychological Corporation.

Wellman, H. M. (1992). *The Child's Theory of Mind.* The MIT Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta - analysis of theory - of - mind development: The truth about false belief. *Child Development, 72*(3), 655-684.

Wellman, H. M., & Liu, D. (2004). Scaling of theory - of - mind tasks. *Child Development, 75*(2), 523-541.

West, R. L. (1996). An application of the prefrontal cortex function theory to cognitive aging. *Psychol Bull., 120*, 272-292.

West, R. L. (2000). In defense of the frontal lobe hypothesis of cognitive aging. *J Int Neuropsychol Soc., 6,* 727-729.

Westfall, J. (2015). PANGEA: Power ANalysis for GEneral Anova designs. Retrieved from http://jakewestfall.org/publications/pangea.pdf.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128.

Witvliet, M., Brendgen, M., Van Lier, P. A., Koot, H. M., & Vitaro, F. (2010). Early adolescent depressive symptoms: Prediction from clique isolation, loneliness, and perceived social acceptance. *Journal of Abnormal Child Psychology, 38*(8), 1045-1056.

Wysocka, J., Golec, K., Haman, M., Wolak, T., Kochański, B., & Pluta, A. (2020). Processing false beliefs in preschool children and adults: Developing a set of custom tasks to test the theory of mind in neuroimaging and behavioral research. *Frontiers in Human Neuroscience, 14.*

Young, L., Dodell-Feder, D., & Saxe, R. (2010). What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. *Neuropsychologia, 48*(9), 2658-2664. doi: 10.1016/j.neuropsychologia.2010.05.012.

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.

Zonneveld, H. I., Pruim, R. H., Bos, D., Vrooman, H. A., Muetzel, R. L., Hofman, A., ... & Vernooij, M. W. (2019). Patterns of functional connectivity in an aging population: The Rotterdam Study. *Neuroimage, 189*, 432-444.

Appendices

Appendix A

Appendix A. List of studies and their associated taxonomy.

Study	Year	Task Description	Taxonomic Label
Bernstein DM, Thornton WL, Sommerville JA.	2011	Sandbox Task	ToM comics / cartoons
Calso C, Besnard J, Allain P.	2019	MPS-TOMQ; ToM Picture Story Task variant	ToM comics / cartoons
Calso C, Besnard J, Allain P.	2019	ToM Picture Story Task	ToM comics / cartoons
Calso C, Besnard J, Allain P.	2020	ToM Picture Story Task	ToM comics / cartoons
Duval C, Piolino P, Bejanin A, Eustache F, Desgranges B.	2011	False Belief: Comic Strips	ToM comics / cartoons
Duval C, Piolino P, Bejanin A, Eustache F, Desgranges B.	2011	ToM Comic Strips - Attribution of intention	ToM comics / cartoons
El Haj M, Raffard S, GÃ©ly-Nargeot MC.	2016	False Belief: Comic Strips	ToM comics / cartoons
Fischer AL, O'Rourke N, Loken Thornton W.	2017	Yoni - cognitive	ToM comics / cartoons
Phillips LH, Bull R, Allen R, Insch P, Burr K, Ogg W.	2011	False Belief: Comic Strips	ToM comics / cartoons
Baksh RA, Bugeja T, MacPherson SE.	2020	ESCoT (Cog ToM Only)	ToM movies
Grainger SA, Steinvik HR, Henry JD, Phillips LH.	2019	TASIT	ToM movies
Johansson Nolaker E, Murray K, HappÃ© F, Charlton RA.	2018	Strange Stories Film Task	ToM movies
Lecce S, Ceccato I, Cavallini E.	2019	MASC	ToM movies
Reiter AMF, Kanske P, Eppinger B, Li SC.	2017	EmpaToM	ToM movies
Zhang X, Lecce S, Ceccato I, Cavallini E, Zhang L, Chen T.	2018	Geometric Shapes Animation	ToM movies
Sarah A. Grainger, Vidhya Rakunathan, Alexandra G. Adams, Allana L. Canty & Julie D. Henry	2020	VAMA (cog.)	ToM-VR Immersive
Cavallini E, Lecce S, Bottiroli S, Palladino P, Pagnin A.	2013	Strange Stories	Vignettes - False belief
Elizabeth A. Maylor, Jane M. Moulson, Ann-Marie Muncer and Louise A. Taylor	2002	Strange Stories	Vignettes - False belief
Elizabeth A. Maylor, Jane M. Moulson, Ann-Marie Muncer and Louise A. Taylor	2002	Strange Stories	Vignettes - False belief
Fischer AL, O'Rourke N, Loken Thornton W.	2017	Strange Stories	Vignettes - False belief
Franco MG, Smith PK.	2013	Strange Stories	Vignettes - False belief

Grainger SA, Henry JD, Naughtin CK, Comino MS, Dux PE.	2018	ToM Stories	Vignettes - False belief
Happé, FG., Winner, E., Brownell, H.	1998	Strange Stories	Vignettes - False belief
Hughes C, Cassidy BS, Faskowitz J, Avena-Koenigsberger A, Sporns O, Krendl AC.	2019	False Belief False Photo	Vignettes - False belief
Lecce S, Ceccato I, Cavallini E.	2019	Strange Stories	Vignettes - False belief
Li X, Wang K, Wang F, Tao Q, Xie Y, Cheng Q.	2013	False Belief: Vignette	Vignettes - False belief
Moran JM, Jolly E, Mitchell JP.	2012	False Belief False Photo	Vignettes - False belief
Rakoczy H, Harder-Kasten A, Sturm L.	2012	Strange Stories	Vignettes - False belief
Rakoczy H, Wandt R, Thomas S, Nowak J, Kunzmann U.	2018	Strange Stories	Vignettes - False belief
Sarah A. Grainger, Vidhya Rakunathan, Alexandra G. Adams, Allana L. Canty & Julie D. Henry	2020	Strange Stories	Vignettes - False belief
Slessor G, Phillips LH, Bull R.	2007	Strange Stories	Vignettes - False belief
Sullivan, S., & Ruffman, T.	2004	Strange Stories	Vignettes - False belief
Bottiroli S, Cavallini E, Ceccato I, Vecchi T, Lecce S.	2016	Faux Pas Test	Vignettes - Faux pas
Li X, Wang K, Wang F, Tao Q, Xie Y, Cheng Q.	2013	Faux Pas Test	Vignettes - Faux pas
Zhang X, Lecce S, Ceccato I, Cavallini E, Zhang L, Chen T.	2018	Faux Pas Test	Vignettes - Faux pas

Appendix B

Appendix B. List of studies precluded from the OA education (i.e., <14 years vs. >14 years) sub-analysis, due to either not reporting education data or reporting education in discrete categories.

Zhang X, Lecce S, Ceccato I, Cavallini E, Zhang L, Chen T (2008)
Franco MG, Smith PK. (2013)
Johansson Nolaker E, Murray K, Happé F, Charlton RA. (2018)
Moran JM, Jolly E, Mitchell JP. (2012)
Reiter AMF, Kanske P, Eppinger B, Li SC. (2017)
Sullivan, S., & Ruffman, T. (2004)
Rakoczy H, Harder-Kasten A, Sturm L. (2012)
Phillips LH, Bull R, Allen R, Insch P, Burr K, Ogg W. (2011)
Sarah A. Grainger, Vidhya Rakunathan, Alexandra G. Adams, Allana L. Canty & Julie D. Henry (2020)

Appendix C

Ethics approval letter for behavioural false belief study (Chapter 3).

Life and Health Sciences Ethics Committee's Decision Letter

To: Dr Charlotte Hartwright

Cc: Samantha Brown

Administrator, Life and Health Sciences Ethics Committee

From: Dr Rebecca Knibb

Chair, Life and Health Sciences Ethics Committee

Date: 8/2/2017

Subject: Project #1065: Differences in Social Cognitive Function Across Healthy Adults

Thank you for your submission. The additional information for the above proposal has been considered by the Chair of the LHS Ethics Committee.

Please see below for details of the decision and the approved documents.

Reviewer's recommendation: Favourable opinion

Please see the tabled list below of approved documents:

Documentation	Version/s	Date	Approved
Participant information sheet	Version 2 as per electronic ethics application #1065	Jan 2017	
Consent form	Version 2 as per electronic ethics application #1065	Jan 2017	
Task_ordering	Version 1 as per electronic ethics application #1065	Dec 2016	
all_advertisements	Version 1 as per electronic ethics application #1065	Dec 2016	
Screening_advisory	Version 1 as per electronic ethics application #1065	Dec 2016	
Response to reviewers	As per electronic ethics application #1065	Jan 2017	

After starting your research please notify the LHS Research Ethics Committee of any of the following:

Substantial amendments. Any amendment should be sent as a Word document, with the amendment highlighted. The amendment request must be accompanied by all amended documents, e.g. protocols, participant information sheets, consent forms etc. Please include a version number and amended date to the file name of any amended documentation (e.g. "Ethics Application #100 Protocol v2 amended 17/02/12.doc").

New Investigators

The end of the study

Please email all notifications and reports to lhs_ethics@aston.ac.uk and quote the original project reference number with all correspondence.

Ethics documents can be downloaded from: http://www.ethics.aston.ac.uk/documents-all. Please note that these documents can ONLY be opened using Mozilla Firefox or the latest Internet Explorer version (IE9).

Statement of Compliance

The Committee is constituted in accordance with the Government Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK. In accord with University Regulation REG/11/203(2), this application was considered to have low potential risk and was reviewed by three appropriately qualified members, including the Chair of the Life and Health Sciences Ethics Committee.

Yours sincerely,

Dr Rebecca Knibb
Chair, LHS Ethics Committee

Appendix D

Ethics approval letter for fMRI false belief study (Chapter 4).

Date: 3rd April 2019

Dr Charlotte Hartwright

(Student: Foyzul Rahman) Life and Health Sciences

Dear Charlotte

Study title:	*Thinking about other people's thoughts in healthy ageing*
REC REF:	Ethics application #1486

Confirmation of Ethical Opinion

On behalf of the Committee, I am pleased to confirm a favourable opinion for the above research based on the basis described in the application form, protocol and supporting documentation listed below.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Participant Information Sheet	1.04	03/04/2019
Consent Form	1.04	03/04/2019
Study measures comprising: • Wide Range Achievement Test 4th Edition (WRAT-4) – screening for impairment in reading. • Autism Quotient (AQ-10) – screening for suspected autism. • Mini Addenbrooke's Cognitive Examination (M-ACE) – dementia screening for older adults. • Edinburgh Handedness Inventory – to determine handedness. • Broad Autism Phenotypes Questionnaire (BAPQ) – captures personality and language characteristics that reflect the phenotypic expression of genetic liability to autism		

With the Committee's best wishes for the success of this project. Yours sincerely

Professor Richard Booth

Acting Chair of the University Research Ethics Committee

Appendix E

Aston Brain Centre MRI Safety & Screening Form

MRI CONSENT & SAFETY INFORMATION FORM
Please answer all questions by ticking YES or NO. If you answer YES to any please give DETAILS below

	DO YOU HAVE	YES	NO	DETAILS
1 *	**A pacemaker/artificial heart valve or stent**			
2 *	**Any electrical devices e.g. cochlear implant (in your ear) nerve/n stimulator, implantable pump or other device anywhere in your l**			
3 *	**Aneurysm clips in your head or hydrocephalus shunt**			
4 *	**False eyes**			
5	Any body piercing, tattoos or permanent eye makeup			
	HAVE YOU EVER HAD			
6 *	**Metal fragments in your eyes or shrapnel in your body**			
7 *	**Any surgery within 6 weeks of appointment**			
8 *	**Any surgery which involves implants to the head or body**			
	DO YOU WEAR			
9	Removable dentures or dental brace with metal			
10	A hearing aid			
11	An artificial limb, surgical support or brace			
12	Foil-backed patches for HRT, Nicotine, pain relief, etc.			
	DO YOU SUFFER FROM			
13	Asthma attacks			
14	Epilepsy or Blackouts			
15	Diabetes or thermoregulatory problems			
	FOR WOMEN OF CHILD BEARING AGE ONLY			
16 *	**Are you or could you possibly be pregnant**			
17 *	**Do you have an IUCD / coil in place or sterilisation clips**			

Name and Address of your Doctor:
…………………………………………………………………………………………………….

Due to the very strong magnetic fields it is VERY IMPORTANT for your own safety and the safety of the MRI Unit staff that you do not go into the Magnet room with any objects which may become dislodged.

PLEASE ENSURE THAT YOU HAVE REMOVED ALL METAL OBJECTS INLUDING. ZIPS, JEWELLERY, HAIR GRIPS, WATCHES, CREDIT CARDS, BELTS, MONEY, MOBILE PHONES AS WELL AS COLOURED CONTACT LENS.

To be signed by visitor, contractor or staff to confirm the information provided on this form is correct

Signature: ………………………………………………………………. Date: ….…………………20

Radiographer Signature: …………………..…………………………………………………. Date: …………………….....20