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

Foyzul Rahman[a], Sabrina Javed[a1], Ian A Apperly[b], Peter C Hansen[b], Carol A Holland[a2], Klaus Kessler[a], Charlotte E Hartwright*[a]

* corresponding author

[a] School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham, UK, B4 7ET

[b] School of Psychology, University of Birmingham, Birmingham, UK, B14 2TT

[1] Present address: Division of Psychiatry, Faculty of Brain Sciences, University College London, UK

[2] Present address: Division of Health Research, Faculty of Health and Medicine, Lancaster University, UK, LA1 4YG

Email addresses:

rahmanf2@aston.ac.uk (F. Rahman); sabrina.javed@ucl.ac.uk (S. Javed); i.a.apperly@bham.ac.uk (I. Apperly); p.c.hansen@bham.ac.uk (P. Hansen); c.a.holland@lancaster.ac.uk (C. Holland); k.kessler@aston.ac.uk (K. Kessler); c.hartwright@aston.ac.uk (C. Hartwright)

Conflict of Interest: None

Abstract

Age-related decline in Theory of Mind (ToM) may be due to waning executive control, which is necessary for resolving conflict when reasoning about others' mental states. We assessed how older (OA; n=50) versus younger adults (YA; n=50) 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 motor response speed best explained the degree of conflict experienced through conflicting 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.

Statement of Relevance

We use our ability to interpret what other people think on a daily basis. Termed 'Theory of Mind' (ToM), this is an essential facilitator of social interaction. Previous studies suggest that ToM proficiency declines in later life, which is associated with poorer psychological wellbeing. However, these earlier studies often confounded changes in ToM with broader age-related changes in 'executive functions' (EF) like attention, memory and inhibitory control. There is general consensus that these EFs decline with age, but this is less clear for ToM, due to how ToM and EF interact. This research examined the link between ToM and EF to disentangle the effects of declining EF from any age-related changes in ToM. We show how prior studies may overestimate the decline of ToM in healthy ageing, because of experimental demands which are not essential for a functioning ToM. This information could inform support for older adults' psychological functioning.

Open Practices Statement

This study was formally preregistered following an Open Science Framework protocol, which is available from https://osf.io/dc8ce. The novel ToM task, de-identified data and a supporting code-book are freely available from https://osf.io/9zmx8/.

1. 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 typically deteriorates with healthy ageing (Salthouse, 2010; 2012), but there is also evidence that cognitive conflict is 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' or CofK (Birch & Bloom, 2004; 2007). To make predictions based on an agent's FB, one must inhibit one's own perspective of the true state of affairs – *I know the object is really in the box* – to adopt the other person's – *they think the object is still in the basket*. 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 incongruence in 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 from the existing literature 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. 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 predict poorer performance in OAs (Hughes et al., 2019). Given that prior research shows differentiation between YA and OA in ToM as a function of EF demands (Bottiroli et al., 2016; German and Hehman, 2006), it is important to better understand whether EF changes impact ToM and how conflict affects ToM processing more broadly and in ageing.

In this preregistered study (protocol: https://osf.io/dc8ce) we investigated the association between conflict and ToM in healthy ageing. Throughout, we refer to specific research questions (RQ*n*) and their corresponding hypotheses (RH*n*), as required for the preregistration protocol. First, we aimed to understand which factors are responsible for variation in the difficulty of reasoning about the beliefs of others (RQ1-4). We hypothesized that *both* a participant's knowledge of reality (RH1) *and* incongruence between the participant's "self-perspective" and an agent's "other-perspective" (RH2) would contribute to difficulty. Further, we hypothesised an interaction between these effects (RH3), reflecting the intuition that our judgements of others' perspectives may be most difficult when the other person not only has a different perspective, but we know that it is wrong. These two aspects are confounded in the classic object transfer FB task, consequently, the present paradigm aims to disentangle these as candidate sources of conflict.

Second, we set out to clarify the contribution of attentional cueing to performance costs on FB tasks (RQ4). It has often been noted that giving the correct answer in common FB paradigms might require 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 typical feature of many FB paradigms though, unlike differences between Self- and Other-perspectives, it 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. We hypothesised that there would be greater cognitive effort associated with holding in mind competing Self vs Other perspectives versus managing alternate locations (Other-Other perspectives; RH4). These first aims were assessed across the entire sample.

Third, studies of ToM in healthy ageing generally suggest reduced ToM capacities in later life, where OA show impairment in false-, but not true-belief reasoning (Phillips et al., 2011). However, the cognitive basis of this impairment is unclear and confounded due to methodological limitations of some ToM tasks. By manipulating psychologically relevant parameters within a single ToM task, we aimed to evaluate which cognitive components of belief-reasoning could explain age-specific deficits in performance (RQ5-6). We expected that ageing would be associated with an overall reduction in performance – slower and more error-prone responses – and that OA would show greater difficulty with self-perspective inhibition, attending and managing conflict from multiple cued locations, and handling incongruence between beliefs – all aspects that are pertinent to FB representation (RH5-6). Lastly, we conducted exploratory analyses to further explore the neuropsychological bases of our results. We used standardised measures of EF alongside age to predict the magnitude of conflict elicited by our ToM manipulations.

2. Materials and Methods

For transparency, any deviations from the preregistration are identified within the text of this manuscript using an S appended by a superscript alphanumeric identifier (e.g., [S1a]) which corresponds to a statement in SOM-R Supplement 1.

2.1 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 YA due to a methodological issue and one OA for scoring beyond the cut-off point in a dementia screening measure. Thus, the final sample comprised 100 participants[S1a]; 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). 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. The majority of the OA (57%) were educated to undergraduate degree level or higher. The study was approved by Aston University's Life & Health Sciences Ethics Committee. All participants gave written informed consent prior to participation.

2.2 Statistical Power

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).

2.3 Design

The study design and analysis plan were preregistered; the registration can be found with the DOI: 10.17605/OSF.IO/DC8CE. All data and project materials can be found with DOI: 10.17605/OSF.IO/9ZMX8.

2.3.1 Theory of Mind abilities: False belief task

The current ToM paradigm was based on Apperly et al. (2011) and Hartwright et al. (2012); it can be downloaded in E-prime format from https://osf.io/9zmx8/ (E-Prime 2.0 software; Psychology Software Tools, Pittsburgh, PA). The current task was non-inferential: participants are explicitly made aware of others' belief states. See Supplement 2 for a discussion on the equivalence of inferential and non-inferential ToM tasks.

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 1. The first factor, termed 'Knowledge of Reality' (KoR)[S1B], 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)[S1B], 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)[S1B], 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 1 HERE>

Table 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. 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.

Each trial comprised five static images, followed with a central fixation mark (see Figure 1, Panel A). 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$_0$) 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 1, Panel B. By varying KoR, OOC and SOC, eight conditions were created: KoR$_0$OOC$_0$SOC$_0$; KoR$_1$OOC$_0$SOC$_0$; KoR$_0$OOC$_1$SOC$_0$; KoR$_0$OOC$_0$SOC$_1$; KoR$_1$OOC$_1$SOC$_0$; KoR$_1$OOC$_0$SOC$_1$; KoR$_0$OOC$_1$SOC$_1$; KoR$_1$OOC$_1$SOC$_1$, 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.

<FIGURE 1 ABOUT HERE>

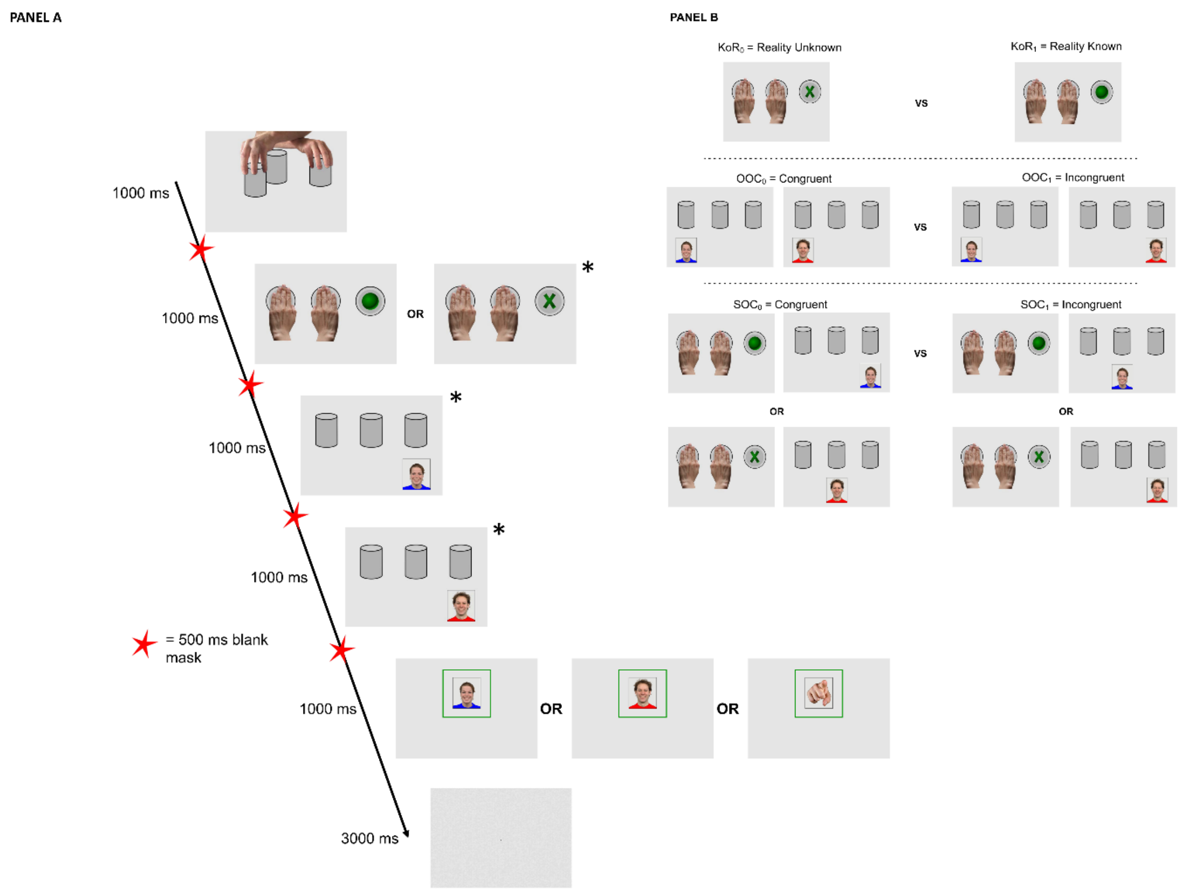


Figure 1. Panel A: A single trial outline with event timings. *Order of events varied across trials. Panel B: Schematic of the three experimental factors and how these were achieved. KoR = Knowledge of Reality: reality unknown (KoR$_0$), where an X was shown to indicate absence of the ball from that 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.

2.3.2 Theory of Mind abilities: – Perspective Taking

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.

2.3.3 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.

2.3.4 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 were excluded from the final sample.

2.4 Procedure

First, participants completed a short training session (detailed in SOM-R Supplement 3) 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 SOM-R Supplement 4 for explicit tests showing no significant cross-group fatigue effects).

2.5 Statistical Analysis

All confirmatory analyses were conducted in SPSS (version 24) and JASP (version 0.12.2). The raw data, summary data and novel test materials used in this study can all be downloaded from https://osf.io/9zmx8/

3. Results

3.1 Sample characteristics

OAs showed poorer performance across all neuropsychological measures. There was no statistically significant difference in self-reported perspective taking (PT) on the IRI (Table 2). Due to equipment failure, no CRT data were acquired for one OA.

<TABLE 2 HERE>

Table 2.

Sample characteristics by group: perspective taking, empathic concern 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 $^{*}p<.05$. $^{**}p\leq.01$. $^{***}p\leq.001$.

3.2 Confirmatory ToM task analyses

3.2.1 False belief task data pre-processing

Prior to statistical analysis, the data were pre-processed as described in the study preregistration (see SOM-R Supplement 5, Table S2 for breakdown of trials removed). 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 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 (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; 404 for YAs and 661 for OAs.

3.2.2 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 2, Panel A) were entered into a four-way mixed ANOVA[S1C] 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 2, Panel B). To be consistent with our preregistered design and analysis protocol, we initially 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, all results are reported in Table 3.

<FIGURE 2 HERE>

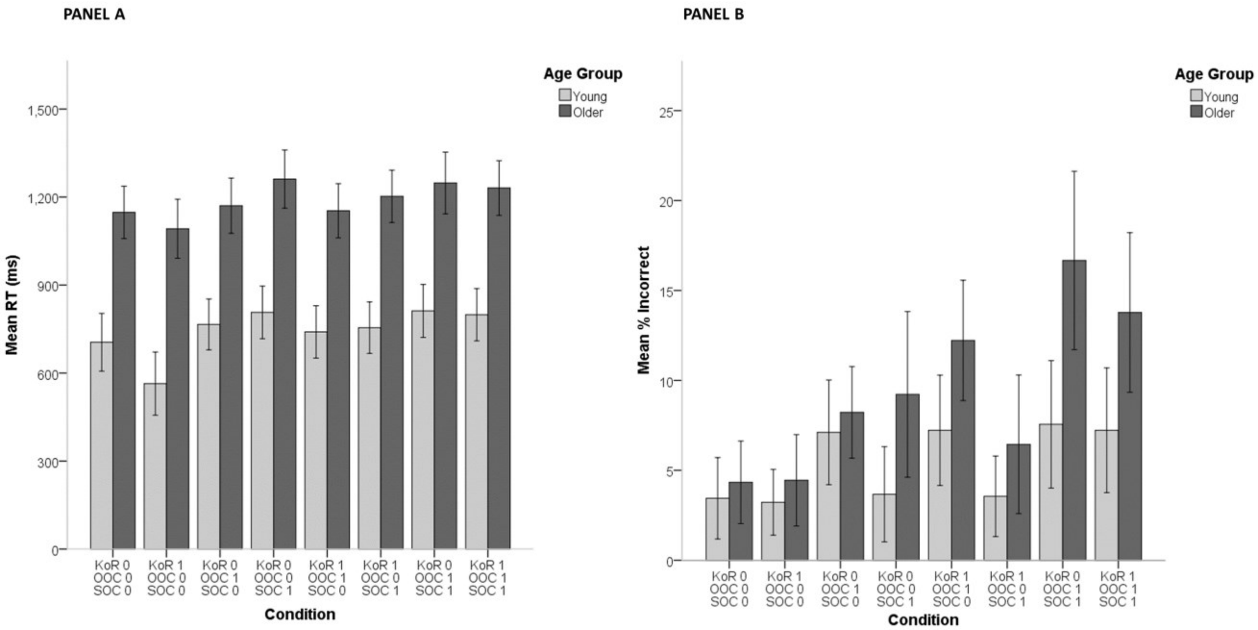


Figure 2. Panel A: Mean condition latencies separated by Group. Panel B: Mean percentage incorrect (accuracy) separated by Group. Error bars: +/- 2 SE.

<TABLE 3 HERE>

Table 3.

Mixed ANOVA results for reaction time and error rate

Effect	ANOVA: Reaction Times (milliseconds)					Sig. Post-hoc Comparisons RT		ANOVA: % Inaccurate Responses					Sig. Post-hoc Comparisons Inaccuracy	
	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 (Young / Older)[RQ5]	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 Knowledge of Reality (KoR)[RQ1]	1	98	47.09	<.001***	.325	KoR$_0$ > KoR$_1$***	47.59	1	98	.39	.536	.325	N/A	-
Main effect of Self-Other Congruence (SOC)[RQ2]	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 Other-Other Congruence (SOC)	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														
Knowledge of Reality (KoR) * Age Group[RQ6]	1	98	2.25	.137	-	N/A	-	1	98	.09	.769	-	N/A	-
Self-Other Congruence (SOC) * Age Group[RQ6]	1	98	.062	.803	-	N/A	-	1	98	5.98	.016*	.058	Older: SOC$_0$ < SOC$_1$***	-4.333
													SOC$_1$: Young < Older*	-6.028
Other-Other Congruence (OOC) * Age Group[RQ6]	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							
Knowledge of Reality (KoR) * Self-Other Congruence (SOC)[RQ3]	1	98	3.74	.056	-	N/A	-	1	98	7.80	.006**	.074	KOR$_0$: SOC$_0$ < SOC$_1$ ***	-3.500
													SOC$_1$: KoR$_0$ > KoR$_1$ *	1.528
Knowledge of Reality (KoR) * Other-Other Congruence (OOC)	1	98	36.05	<.001***	.269	KOR$_0$: OOC$_0$ < OOC$_1$*	-18.83	1	98	2.10	.150	-	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							
Self-Other Congruence (SOC) * Other-Other Congruence (OOC)	1	98	20.07	<.001***	.170	SOC$_0$: OOC$_0$ < OOC$_1$***	-80.29	1	98	1.22	.271	-	N/A	-
						OOC$_0$: SOC$_0$ < SOC$_1$***	-129.03							
						OOC$_1$: SOC$_0$ < SOC$_1$***	-64.89							
Knowledge of Reality (KoR) * Self-Other Congruence (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$_1$***	-6.667
						Young KOR$_1$: SOC$_0$ < SOC$_1$***	-124.52						Older SOC$_0$: KOR$_0$ < KOR$_1$*	-2.055
						Young SOC$_0$: KOR$_0$ > KOR$_1$***	83.33						Older SOC$_1$: KOR$_0$ > KOR$_1$*	2.834
						Young SOC$_1$: KOR$_0$ > KOR$_1$***	32.67						see Table 5	
						See Table 4								
Knowledge of Reality (KoR) * Other-Other Congruence (OOC) * Age Group	1	98	3.62	.060	-	N/A		1	98	1.87	.175	-	N/A	-
Self-Other Congruence (SOC) * Other-Other Congruence (OOC) * Age Group	1	98	4.32	.040*	.042	Young OOC$_0$: SOC$_0$ < SOC$_1$***	-146.14					-	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 6								
Knowledge of Reality (KoR) * Other-Other Congruence (OOC) * Self-Other Congruence (SOC)	1	98	1.83	.180	-	N/A	-	1	98	2.36	.127	-	N/A	-
Knowledge of Reality (KoR) * Other-Other Congruence (OOC) * Self-Other Congruence (SOC) * Group	1	98	2.10	.150	-	N/A	-	1	98	2.35	.248	-	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. RQ: Specific hypotheses were preregistered for these results, where each analysis (RQ1, RQ2 etc) is discussed further within the formal study preregistration; RQn 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.

There was a statistically significant within-subjects (w/s) main effect of KoR in RT where, contrary to our preregistered hypothesis, $KoR_0 > KoR_1$ (RH1); participants were slower when they did not know the location of the ball, although there was no significant effect of KoR in accuracy. The w/s main effect of SOC was significant for RT and accuracy, where, in line with our registered hypothesis, participants were slower and made more errors when Self-Other perspectives where incongruent (SOC_1) versus congruent (SOC_0; RH2). The w/s main effect of OOC was also significant, where incongruence between Other-Other (OOC_1) perspectives was associated with increased costs in RT and accuracy.

Our hypothesis predicting a two-way interaction between SOC and KOR (RH3) was not supported - there was no significant increase in RT/inaccuracy when participants held explicit knowledge of reality (KoR_1) and when their perspectives differed from the target agent's (SOC_1) - a CofK as in the classic FB task. There was a significant interaction between KoR and SOC for accuracy when Self-Other knowledge states were incongruent (SOC_1). A simple effects analysis showed that, unexpectedly, this effect occurred when the participant did not know where the ball was, indicating that participants made more mistakes specifically when reality was unknown (KOR_0) and Self-Other beliefs differed (SOC_1); however, note that this two-way interaction is 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 (see later results and Table 5).

A between-subjects effect of Age Group indicated that OA were significantly slower and more error-prone than YA, as we predicted (RH5). We further expected an interaction between Age Group and each of the w/s factors, KoR, SOC and OOC (RH6). 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. Nonetheless, there was an interaction between SOC and Age Group in accuracy, consistent with our preregistered hypothesis. As detailed in Table 3, simple effects analyses demonstrated that only OA made significantly more mistakes when their knowledge was incongruent (SOC_1) versus congruent (SOC_0) with the target agent and OA were significantly more error-prone than YA when Self-Other knowledge was incongruent (SOC_1).

There were also several interaction effects which we had not registered any predictions for and, as such, should be treated as exploratory; however, these are subsequently reported for transparency.

There was a significant two-way interaction between KoR and OOC in RT. Simple effects analyses indicated that the effect of OOC was such that participants were consistently slower to respond when the two agents' beliefs were incongruent (OOC_1) versus congruent (OOC_0), and manipulating KOR consistently resulted in slower responses when reality was unknown (KOR_0) versus known (KOR_1). Further, there was an advantage to RTs when the two agents' beliefs were congruent (OOC_0) and the ball's location known (KOR_1): when the two agents' beliefs differed (OOC_1), participants slowed significantly more for both reality unknown (KOR_0) and known (KOR_1).

There was also a significant two-way interaction between SOC and OOC in RT. Here, the effect of congruency of Other-Other perspectives ($OOC_1 > OOC_0$) was only significant when Self-Other perspectives were congruent (SOC_0).

There was a three-way interaction between KOR, SOC and Age Group in RT and in accuracy. These interactions were interrogated using a further four, two-way repeated-measures ANOVAs, one for each age group separated by RT and accuracy. For RT, as shown in Table 4, there was a significant interaction between KoR on SOC in YAs ($F(1, 49) = 15.45$, $p < .001$, ηp2= .240), but not OAs ($F(1, 49) = .00$, $p = .949$, ηp2= .000). For YAs, while always slower when Self-Other perspectives were incongruent (SOC_1), the effect of KoR was larger when conflict between Self- and Other-perspectives was absent (SOC_0) versus present (SOC_1) and, again in YAs, the effect of SOC was larger when KoR was known (KOR_1) versus unknown (KOR_0). Contrary to the RT data, the two-way interaction in accuracy between KoR and SOC was non-significant in YAs ($F(1, 49) = .04$, $p = .836$, ηp2= .001) but significant in OAs ($F(1, 49) = 9.07$, $p = .004$, ηp2= .156). Table 5 shows that in OAs, there was no effect of SOC on accuracy when reality was known (KoR_1). When reality was unknown (KoR_0), incongruent perspectives between the participant and target agent (SOC_1) caused OAs to become more error prone than when their perspectives were congruent (SOC_0). There was also an effect of KoR when Self-Other perspectives were congruent (SOC_0), where knowing about reality (KOR_1) increased error-rate. Conversely, when Self-Other perspectives were incongruent (SOC_1), not knowing about reality (KOR_0) increased accuracy versus knowing (KoR_1).

<TABLE 4 HERE>

Table 4.

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

	$\bar{X}_0$	$\bar{X}_1$	$\bar{X}$ difference ($_{0-1}$)	*SE*	P_{bonf}
	SOC				
KoR_0	735.451	809.320	-73.870	10.633	< .001***
KoR_1	652.126	776.650	-124.524	9.758	< .001***
	KoR				
SOC_0	735.451	652.126	83.325	11.786	< .001***
SOC_1	809.320	776.650	32.670	7.987	< .001***

Note. *p<.05. **p<u><</u>.01. ***p<u><</u>.001.

<TABLE 5 HERE>

Table 5.

*Pairwise comparisons of percent incorrect in OAs for KoR * SOC*

	$\bar{X}_0$	$\bar{X}_1$	$\bar{X}$ difference ($_{0-1}$)	*SE*	P_{bonf}
	SOC				
KoR_0	6.279	12.946	-6.667	1.848	0.001***
KoR_1	8.334	10.112	-1.77	1.512	.245
	KoR				
SOC_0	6.279	8.334	-2.055	0.839	0.018*
SOC_1	12.946	10.112	2.834	1.221	0.024*

Note. *p<.05. **p<u><</u>.01. ***p<u><</u>.001.

There was also a three-way interaction between Age, OOC and SOC in RT, ($F(1, 98) = 4.32$, p = .040, ηp2 = .042). Two separate repeated-measures ANOVAs indicated the interaction between OOC and SOC was statistically significant in YA ($F(1, 49) = 41.38$, p < .001, ηp2= .458) but not OA ($F(1, 49) = 1.95$, p = .169). In both age groups, a conflicting (SOC_1) versus a congruent Self-Other perspective (SOC_0) resulted in slower RTs overall ($SOC_1 > SOC_0$). Similarly, conflicting versus congruent perspectives between two agents also resulted in slower RTs in both groups ($OOC_1 > OOC_0$). YAs demonstrated a greater effect of SOC when the two agents' perspectives were congruent (OOC_0) versus incongruent (OOC_1). Likewise, the congruency of the two agents' perspectives had a much larger effect on RT when Self-Other perspectives were aligned (SOC_0); Table 6.

<TABLE 6 HERE>

Table 6.

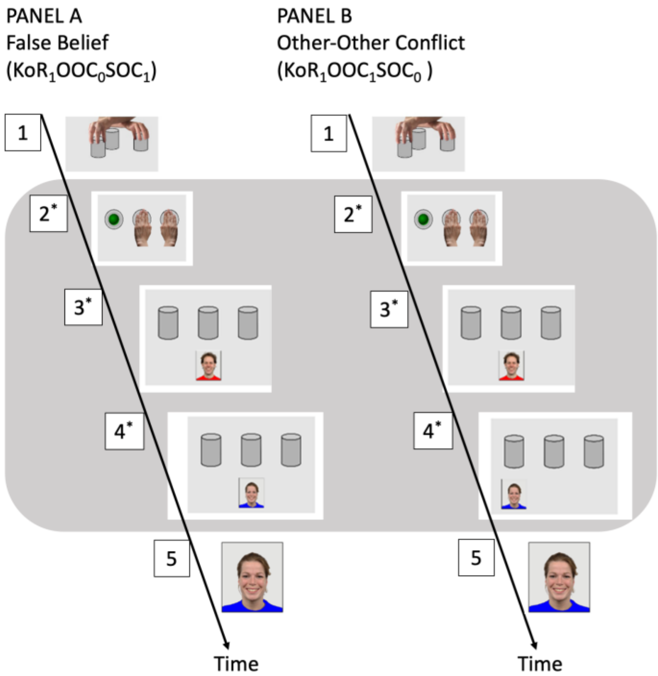
*Pairwise comparisons of RT in YAs for OOC * SOC*

	$\bar{X}_0$	$\bar{X}_1$	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.

Responding correctly in a FB paradigm might require a participant to shift attention between competing locations, for example, where the target object is located and where the target agent believes it is located. This is a typical feature of FB paradigms though, unlike incongruence between Self- and Other-perspectives, it is not an essential feature of FB problems. On this basis, for RH4, 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). Assessing this required us to compare data from two pre-specified conditions within our ToM task: a classic FB condition, where two locations are cued – where the ball really is, and where the agent falsely believes the ball is – creating incongruence in Self-Other perspectives, and a condition like the original FB task where two competing locations are cued but, unlike the original FB, the target agent's and participant's beliefs are congruent. We illustrate these two conditions in Figure 3. 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.

<FIGURE 3 HERE>



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. 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). Each trial exemplar in Panels A and B is indicative of the order of 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.

3.3 Additional Exploratory Analyses

3.3.1 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, YA 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 Table 4, YAs were slowest overall when reasoning about an agent with a FB when reality was unknown (KOR_0). Moreover, OA showed the same pattern in the accuracy of their responses (Table 5); if they knew where the ball was, and the agent's belief was false, they were more error-prone than if the agent's belief was true, but OAs made the most errors 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 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).

3.3.2 ToM, Ageing and EF

Correlation analyses 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). All correlation coefficients are reported in SOM-R Supplement 6.

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 comprised 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, as reported in Table 2, SST, AST, SWM and CRT (Table 7).

<TABLE 7 HERE>

Table 7

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 CANTAB	.264	2.761	.007**
																	IRI PT	-.231	-2.413	.018*
SOC	1	97	5.217	.025*	.051	0.041	AST CANTAB	.226	2.284	.025*	1	96	9.023	<.001***	.158	.141	CRT CANTAB	.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. ns - non-significant: no statistically significant predictive models identified. *$p<.05$. **$p\leq.01$. ***$p\leq.001$.

Table 7 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 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 introduced by varying Self-Other perspectives.

4. Discussion

4.1 What factors affect variation in the difficulty of reasoning about the beliefs of others? (RQ1-4)

We hypothesised that KoR could interfere with one's ability to reason (RH1), due to bias towards own self-perspective – a CofK (Birch & Bloom, 2004; 2007); however, this was not realised in the way predicted. Participants were *faster* when they knew reality. Compatible with our predictions (RH3), the effect of KoR was influenced by congruence of Self-Other perspectives. Incongruent Self-Other beliefs resulted in costs as predicted (RH2), and, as outlined in RH3, explicit KoR caused interference when representing incongruent Self-Other beliefs. Still, greater interference resulted from not knowing reality, particularly when the target agent held a FB.

To explore the unexpected finding for KoR, we considered studies of belief-desire reasoning, where both the agent's belief- (true/false) and desire-states (approach/avoid) are manipulated. Here, an agent's desire-state is varied via a given preference to either seek out, or to avoid, a target object, for example (see Apperly et al., 2011; German and Hehman, 2006; Hartwright et al., 2012; Leslie et al., 2005). This work shows that people exhibit increased difficulty processing false- versus true-beliefs and avoid- versus approach-desires, where a FB combined with an avoidance desire attracts maximal processing costs. Coupled with prior research showing that we anticipate others' behaviour to fulfil, rather than conflict, their desires (Ferguson and Breheny, 2011), this suggests that it was implicit in our task that the agent wanted to avoid the empty location. Consequently, knowledge of the empty location, though undisclosed to the agent, reflects an avoidance desire, irrespective of that agent's belief. This supposition implies that any condition highlighting an empty location would be effortful, and that mentalizing about an agent with a FB regarding the empty location would be doubly effortful: participants must inhibit their knowledge that the agent must avoid the location they (falsely) believe to be true. To test this, we compared data from a classic FB condition versus our novel, reality unknown FB condition. Exploratory analyses supported our assertion: participants took longer to resolve a reality unknown FB, which suggests that KoR itself is not the cause of the CofK. We propose that participants' initial internal reference towards an agent's desire can create conflict, in our case, resulting in re-direction away from the empty cup. This indicates that the salience of self-knowledge relates to the agent's higher order intentions, which determines the CofK.

Additionally, we proposed that holding in mind a competing Self-Other perspective – a FB – would require greater effort than when shifting attention between competing locations (RH4); the latter a common, but not requisite, feature of FB problems. We compared two conditions where the number of locations cued was identical, but only one explicitly required representing a FB. Participants were slower but, contrary to our predictions, more accurate, when representing FBs versus managing competing locations. This suggests that representing a belief that differs from one's own may recruit different processes to those for managing competing attentional cues, which may not be more effortful in itself. Indeed, the accuracy data suggest greater proficiency with inhibiting a conflicting self-perspective compared with inhibiting an irrelevant spatial cue. This finding 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). This supposition is further supported by differing predictors for SOC and OOC within our exploratory regression analyses, discussed shortly.

4.2 Which cognitive components of belief-reasoning explain age-specific differences in ToM? (RQ5-6)

We hypothesised that OAs would demonstrate poorer performance overall compared to YAs (RH5), and that each experimental factor would interact with age (RH6). Consistent with German and Hehman (2006) and Hughes et al. (2019), OAs were generally slower and less accurate than YA. However, our data suggest that poorer performance in OA may be largely attributable to particular experimental features – specifically the need to switch attention between competing cued locations – which disproportionately impact OAs. OA suffered greater costs to both speed and accuracy than YA when managing conflicting cued locations (OOC). Age Group was the only significant predictor of the magnitude of cost to RT in OOC, whereas individual differences in spatial working memory and self-reported ToM proficiency predicted the cost to error-rate. Thus, while the cost of incongruent Other-Other perspectives to error-rate was best explained by how well participants could retain and manipulate visuospatial information, age was relevant to RT costs when managing competing locations. Conversely, both age groups' RTs were similarly affected by manipulating SOC, our core ToM manipulation: they were comparably slower when representing an incongruent, rather than congruent, Self-Other perspective, although OA were more error-prone than YA specifically for conflicting Self-Other perspectives. Age group did not predict the magnitude of conflict elicited by varying congruence of perspectives in RT or accuracy. Instead, attentional capacity was most predictive of uncertainty introduced into RT, and Motor-response speed and self-reported ToM proficiency for the cost of incongruence to error-rate. These data suggest that individual differences in attentional capacity might best explain RT performance when managing competing Self-Other perspectives, and that errors may reflect limitations in EF and motor-response speed rather than an age-related decline in ToM proficiency per se. Moreover, prior studies may inflate age-related changes in ToM capacity, due to incidental task demands which disadvantage OAs.

5. 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 attentional demands in future studies of ageing and ToM.

Author contributions

FR, CEH, KK, IAA, PCH and CAH developed the research concept. FR and CEH designed the experiment. FR and SJ collected the data. FR, CEH and SJ analysed the data. FR and CEH wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Thank you Evelyn Egyir, Elizabeth Essex and Maria Nazakat for assisting with data collection.

References

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.

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.

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.

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.

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.

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

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

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.

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.

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.

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.

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.

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.

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

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.

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.

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.

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.

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.

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.

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

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.