The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 2025, **80**(1), gbae188
https://doi.org/10.1093/geronb/gbae188
Advance access publication: 13 November 2024
Research Article

Age Differences in Mental State Inference of Sarcasm: Contributions of Facial Emotion Recognition and Cognitive Performance

Colleen Hughes, PhD,[1,*] Didem Pehlivanoglu, PhD,[2] Amber Heemskerk, BSc,[3] Rebecca Polk, PhD,[2] Gary R. Turner, PhD,[4] Natalie C. Ebner, PhD,[2,5] and R. Nathan Spreng, PhD[6]

[1]Psychological and Brain Sciences Department, Indiana University, Bloomington, Indiana, USA.
[2]Department of Psychology, University of Florida, Gainesville, Florida, USA.
[3]School of Interdisciplinary Forensics, Arizona State University, Phoenix, Arizona, USA.
[4]Department of Psychology, York University, Toronto, Ontario, Canada.
[5]Center for Cognitive Aging and Memory, McKnight Brain Institute, University of Florida, Gainesville, Florida, USA.
[6]Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada.

*Address correspondence to: Colleen Hughes, PhD. E-mail: collhugh@iu.edu

Decision Editor: Martina Luchetti, PhD (Psychological Sciences Section)

Abstract

Objectives: Older age is associated with poorer ability to accurately infer mental states, but some mental states are more complex than others. Sarcasm is a complex mental state because the literal and intended meaning of a speaker's words are in opposition. Individuals must rely on additional cues (e.g., facial expressions, intonation) for accurate inference. We hypothesized that understanding of sarcastic versus sincere exchanges would be more sensitive to age-related difficulty in mental state understanding.

Methods: We examined accuracy at identifying sarcasm among 263 adults (ages 18–90 years) using videos of social interactions in The Awareness of Social Inference Test. Hypotheses were tested using a logistic linear mixed effects model predicting correct/incorrect trial-level responses. To characterize why sarcasm differed with age, we measured 2 abilities commonly implicated in mental state understanding: facial emotion recognition and cognitive performance.

Results: Sarcasm understanding declined with age, whereas understanding of sincere exchanges did not. Both better emotion recognition and cognitive performance related to better understanding of sarcastic but not sincere exchanges. Only cognitive performance showed an age-related effect such that the cognitive performance among the oldest participants facilitated their understanding of both sarcastic and sincere exchanges.

Discussion: We showed that individual variation related to age and social and cognitive performance is more pronounced when the use of multiple mental state cues is more (sarcasm) or less (sincerity) necessary for accurate understanding of social interactions. Naturalistic paradigms involving multiple mental state cues can address important questions about how older adults make decisions in the real world.

Keywords: Experimental behavioral analysis, Mentalizing, Social cognitive aging, Theory of mind

Understanding others' thoughts, feelings, and intentions is a fundamental aspect of human social behavior (Baumeister & Leary, 1995). The ability to infer such mental states facilitates positive social interactions that may bolster well-being and health against the pernicious effects of loneliness (Cacioppo et al., 2014). This ability becomes especially important in aging given that older adults' social networks emphasize relationship quality versus quantity (Cornwell & Waite, 2009), perhaps to maximize the positivity of social interactions (English & Carstensen, 2014). Concurrent changes in socioemotional and cognitive processes in older adulthood may explain declines in accurate inferences about others' mental states (Hamilton et al., 2022). However, there are multiple ways to identify the mental states of others, and often the multiplicity of such cues is not present in experimental contexts (Henry et al., 2013). It remains an open question if or why older adults may have less accurate mental state inference compared with younger adults in more naturalistic paradigms where multiple cues are available. In the current work, we examined mental state inference from dynamic videos of social interactions among participants aged 18–90 years to characterize this process across the adult lifespan. Moreover, we tested the extent to which facial emotion recognition and cognitive performance—two core abilities that decline in older adulthood (Park & Reuter-Lorenz, 2009; Ruffman et al., 2008)—contribute to mental state inference.

Difficulty at accurately identifying mental states may arise from the type of information available in a scenario. Past investigations largely focused on mental state inference using single modalities (e.g., verbal: written stories [Zaitchik,

Received: May 7 2024; Editorial Decision Date: October 20 2024.

1990]; visual: static images of facial expressions of basic emotions [Ekman, 2006]) and found moderate age-related difficulties (Cassidy et al., 2021; Ebner & Johnson, 2009; Henry et al., 2013; Hughes et al., 2019; Riediger et al., 2011). Facial expressions represent one channel of information about what someone else is thinking or feeling (Henry et al., 2013). A widely used paradigm for assessing facial emotion recognition is the Reading the Mind in the Eyes Test (RMET; Baron-Cohen et al., 2001). In the RMET, participants are presented with static, black and white images of facial expressions of complex emotions (e.g., arrogant, playful) cropped to the eye area, which limits ceiling effects (Baron-Cohen et al., 2001), thereby making it useful for assessing individual variation in normative aging. The RMET and comparable paradigms consistently reveal lower accuracy at identifying facial emotions among older than younger adults (Cassidy et al., 2021; Ebner & Johnson, 2009; Henry et al., 2013; Riediger et al., 2011). Emerging work, however, suggests that age differences are less pronounced when emotion recognition involves dynamic stimuli (e.g., faces morphing from neutral to positive expressions; Hayes et al., 2020; Holland et al., 2019). This finding raises the possibility that older adults may benefit more than younger adults from information conveyed in dynamic aspects of an unfolding expression or social interaction.

An important factor influencing mental state understanding is the multiplicity of cues available for inference (e.g., having faces, bodies, words, and their intonation). Paradigms involving multiple cues may aim to simulate naturalistic social interactions, which could reduce or eliminate age differences as seen for dynamic stimuli within a single cue paradigm (Hayes et al., 2020; Holland et al., 2019). For example, individuals can reach a correct understanding of a speaker's meaning by focusing on only one type of cue or compensating for difficulty in one modality (e.g., facial emotion recognition) by competence in another (e.g., vocal emotion recognition). Extant evidence suggests, however, that older adults have worse accuracy at inferring mental states when multiple types of cues are presented using dynamic stimuli such as audiovisual clips of emotion portrayals (Polk et al., 2022) and social interactions from a TV show (Krendl et al., 2022). These findings could be reconciled under the premise that having multiple cues to mental state inference can, in some cases, increase complexity. The example previously may reflect successful strategies for understanding a sincere or literal exchange (i.e., the meaning of the speaker's words is congruous with their intended meaning [McDonald et al., 2003]). Not all social exchanges are literal or sincere, however. Sarcasm, for instance, is ironic in that the speaker intends that the listener detect that the speaker's literal and intended meaning are in opposition (Shamay-Tsoory et al., 2005; Winner & Leekam, 1991). Because the literal meaning of the speaker's words cannot be used in isolation for accurate understanding, sarcasm reflects a more complex mental state than sincerity.

A previous study found age-related difficulty for the detection of sarcasm but not sincerity (Phillips et al., 2015). Older versus younger adults may exhibit worse complex mental state understanding because of increased reliance on social and cognitive abilities that decline with age (Gurera & Isaacowitz, 2019; Park & Reuter-Lorenz, 2009; Ruffman et al., 2008), possibly due to age-related changes in brain structure and function subserving these capacities (Spreng & Turner, 2019). For example, older adults may rely more on good cognitive function to hold in working memory both the speaker's words and accurate understanding of their intonation and/or to reconcile discordant interpretations from failing to accurately interpret some cues but not others (Wang & Su, 2013). Also, sarcasm can serve positive and negative social relational functions, such as increasing feelings of social connectedness via humor (Huang & Galinsky, 2023) or delivering criticism (Shamay-Tsoory et al., 2005). Interpreting the speaker's mental state informs the valence of the interaction, with likely particular importance in aging because older adults are motivated to prefer and selectively engage in emotional positivity (Hess, 2014; Reed et al., 2014). By contrasting sarcastic and sincere exchanges, we gain a nuanced theoretical picture of how older adults do or do not successfully use and integrate mental state cues when they are more (sarcasm) or less (sincerity) necessary for accurate understanding.

We investigated the association between age and participants' understanding of sarcastic versus sincere social interactions via audiovisual stimuli from The Awareness of Social Inference Test (TASIT; McDonald et al., 2003). The TASIT was originally developed to index clinical-level social impairment in traumatic brain injury. Prior studies mostly collapsed across sincere and sarcastic exchanges as patients showed difficulty with both (Burdon et al., 2016; Grainger et al., 2018; Lavrencic et al., 2016). Capturing sensitivity to individual variation in healthy adults, a previous study validated that older adults were worse than younger- and middle-aged adults at detecting sarcasm from TASIT videos (Phillips et al., 2015). We anticipated replicating that older age was associated with worse understanding of sarcastic, but not sincere, exchanges (Hypothesis 1). If supported, this finding would reflect specific impairment in complex mental state understanding despite the possible advantages of having multiple and dynamic social cues for both exchange types.

Age-related difficulty in decoding more complex mental states may emerge as the result of deficits in specific social skills (e.g., facial emotion recognition) and/or cognitive functioning to integrate multiple cues. Facial emotion recognition mediated age deficits in accurate understanding of sarcastic exchanges on the TASIT (Phillips et al., 2015). What has not been previously tested is whether facial emotion recognition is uniquely helpful for sarcastic versus sincere exchanges. This effect would indicate whether individual variation in facial emotion recognition affects more versus less complex mental state understanding. Here, we measured facial emotion recognition using the RMET because it is sensitive to variability in cognitively healthy aging and it evaluates complex facial emotions that may be more relevant to naturalistic social interactions. Regarding cognitive contributions, the literature linking specific cognitive abilities with age-related deficits in mental state understanding is heterogenous (Charlton et al., 2009; Fernandes et al., 2021; Fischer et al., 2017; Moran, 2013). For sarcasm understanding, most studies examine specific cognitive abilities in isolation (e.g., working memory [Phillips et al., 2015]; working memory updating versus inhibition [Phillips et al., 2011]) or among clinical cohorts (e.g., traumatic brain injury; McDonald et al., 2006). Given extant heterogeneity, we measured cognitive performance using a validated composite score of executive abilities (Lachman et al., 2014). We hypothesized that poorer understanding of sarcasm but not sincerity with greater age would be, at least in part, related to older adults' worse facial emotion recognition (Hypothesis 2A) and/or worse cognitive performance (Hypothesis 2B).

Method

Transparency and Openness

Data for this analysis were taken from a larger study on social-cognitive aging, and only measures relevant to this analysis were reported here. All analyses were conducted with the statistical package R v.4.2.1. We used the lme4 package for mixed effects models; model p-values were calculated using *lmerTest* with statistical significance set to $p < .05$; confidence intervals (CI) were calculated via the base R function *confint*; and the emmeans package was used to calculate pairwise contrasts to characterize interaction effects. Data and code supporting the analysis are presented at https://osf.io/td3eg/. Given the secondary data-analytic nature of this research, the study design, hypotheses, and analytic plan were not preregistered.

Participants

The University of Florida Institutional Review Board approved this study. All participants provided written informed consent. Power analyses using the R package WebPower indicated approximately 200 participants were needed to detect a moderate interaction between age group and exchange type ($f = 0.23$; based on $\eta_p^2 = 0.05$ for the same interaction in Phillips et al., 2015) with at least 80% power and $\alpha = 0.05$. Participants were recruited from the local community using fliers, handouts, mailouts, and advertisements on websites for research studies (e.g., ResearchMatch, newspaper and radio advertisements, word-of-mouth, an established laboratory participant pool, and Healthstreet, a university-affiliated community recruitment and outreach program). All participants received monetary compensation for their participation.

Only participants with complete TASIT data were included in these analyses. The analyzed sample comprised 263 participants ($M_{age} = 45.48$ years, $SD = 21.82$, range 18–90; 78% women, 22% men; 74% White, 8% Black, 6% Asian, 0.4% American Indian, 5% multiracial, 6% not reported; 13% Hispanic ethnicity, 86% non-Hispanic, 0.8% not reported; $M_{yearsedu} = 15.91$, $SD = 2.77$, 16 not reported). Participants with missing scores on the emotion recognition ($n = 20$) and cognition ($n = 5$) measures were excluded from models that included those variables. Participants with ($n = 238$) and without these scores ($n = 25$) did not differ in age or number of correct responses on the TASIT, ts < 1.37, ps > 0.18. Analyses imputing missing scores are reported in Supplementary Material and Supplementary Table 6, and the conclusions remain unchanged after imputation.

Materials

The Awareness of Social Inference Test

The TASIT was originally developed as a clinical assessment tool for social perception (McDonald, 2012; McDonald et al., 2003). Relevant here, TASIT-Part 2 assesses participants' accuracy at understanding others' mental states (e.g., thoughts and feelings) from audiovisual stimuli of social interactions portrayed by professional actors. Participants watched 15 videos (approx. 30 s each) and after each answered four probe questions about their understanding of the speakers' actions, statements, thoughts, and feelings. All probe questions (e.g., "Is Ruth trying to pressure Gary into helping her?") had the following response options: "Yes," "No," or "Don't Know." The videos featured Australian male and female actors portraying scripted content. Scores were similar between North American and Australian samples, genders, and across varying levels of education (McDonald et al., 2018). The videos varied whether the speaker's words were meant to be sincere (literal; five videos) or sarcastic (nonliteral). The sarcasm videos portrayed either simple sarcasm (i.e., potentially sincere words—*"You have been a great help!"*—spoken sarcastically; five videos) or paradoxical sarcasm (i.e., the words could only make sense if everyone in the interaction understood the speaker was being sarcastic—*"Do you have your passport? No I tore it up. Okay good!"*; five videos). We did not hypothesize differences based on the type of sarcasm because there is little evidence that cognitively healthy adults find one type more difficult than the other (McDonald et al., 2003; Phillips et al., 2015); however, we analyzed them separately for comprehensiveness and transparency. Participants' response to each probe item for each video was marked as correct (1) or incorrect (0) using the provided scoring manual (McDonald et al., n.d.). We administered both two task versions, counterbalanced across participants (Form A: $n = 138$; Form B: $n = 125$). Each form had the same number and order of items by exchange type but varied the scenarios among sincere and simple sarcasm items (e.g., Form A might have a speaker being sincere about a promotion, but Form B has the speaker being [simple] sarcastic about a promotion). The paradoxical sarcasm items had different scenarios in each version. TASIT-Part 2 had a moderate to good test–retest reliability of $r = 0.62$–0.88 in a prior study (McDonald et al., 2006). Further, construct validity was good as reflected by a strong association with second-order, but not first-order, theory of mind (more and less complex mental state understanding, respectively); for a review, see McDonald, 2012.

Reading the Mind in the Eyes Test

To measure emotion recognition from faces, we administered a computerized version of the RMET (Baron-Cohen et al., 2001). Participants progressed through 36 trials at their own pace. On each trial, participants saw an image of a face cropped to the person's eyes in the center of the screen and four trait words, different sets for each trial, around the corners of the image. Participants chose which word best reflected what the person in the image was feeling. A dictionary for the trait words was provided upon request. The number of correct responses was totaled and standardized for regression analysis, where higher values indicated better emotion recognition performance. Internal consistency across items, after classifying responses as accurate or inaccurate, was similar to prior reports, albeit low (KR-20 [ranging 0–1 where values closer to 1 indicate high reliability] = 0.68 [Higgins et al., 2023]).

Brief Test of Adult Cognition by Telephone

We administered the Brief Test of Adult Cognition by Telephone (BTACT) by video call. The BTACT is a 30-min test battery, sensitive to cognitive status in healthy aging and community-based samples without ceiling effects, and validated as having convergence with in-lab assessments (Lachman et al., 2014). The BTACT comprises the following subtests: (a) total number of digits correctly produced minus the number of errors in the *backward counting* subtest measured speed of processing; (b) highest number of digits reached in the *digit backward recall* subtest measured working memory updating; (c) total number of unique correct responses in the *category fluency* subtest measured verbal fluency; (d) total number of correct items in the *number* *series* subtest measured inductive reasoning; (e) total number of correct items on the cued switch trials in the *stop and go* subtest measured performance to flexibly shift attention; and (f) total number of unique correct responses in the *immediate and delayed word list recall* measured short- and long-term episodic verbal memory. A prior confirmatory factor analysis showed that a two-factor solution (executive function and memory) was a good fit; therefore, we calculated a composite score of cognitive performance from the five non-memory subtests by standardizing each and then averaging them together (Lachman et al., 2014), with higher values indicating better cognitive performance.

Participants were first screened for study eligibility over the telephone. Then, we administered the BTACT and TASIT during a video call, and participants completed the RMET and demographic questions via REDCap—an online platform for data collection.

Analytic Approach

The TASIT is commonly scored as a total of correct responses to an index clinical deficit in social cognition, for which the measure was originally developed (McDonald et al., 2003). For comparison to prior work, bivariate correlations between the total number of correct responses for each exchange type (sincere, simple sarcasm, paradoxical sarcasm), age, facial emotion recognition performance, and cognitive performance are reported in Supplementary Table 1. Here, we analyzed the data using a logistic linear mixed effects model to serve our goal of having sensitivity to individual variation in a cognitively healthy adult lifespan sample by partialing out random variance associated with the nested task structure (i.e., probe questions within video and videos within exchange types). Correspondingly, the outcome of all analyses was trial-level accuracy (0 = incorrect, 1 = correct) on the TASIT, and we included random intercepts of participant, video, and probe question type. To test Hypothesis 1, standardized chronological age, exchange type (sincere, simple sarcasm, paradoxical sarcasm), and their interaction served as fixed effects. Model fit as measured by AIC was not significantly improved for any models when including a fixed effect of gender or years of education as covariates. We then separately tested Hypothesis 2A (regarding facial emotion recognition) and Hypothesis 2B (regarding cognitive performance) by adding their respective main effects and interactions with age and exchange type as fixed effects in the previously-described logistic linear mixed effects model. Finally, we tested a model including both facial emotion recognition and cognitive performance and their interactions to delineate their unique contributions to TASIT accuracy. Post hoc analyses to decompose main effects and interactions were conducted without a p-value adjustment to fully characterize the effects.

Results

Older Age Related to Less Accurate Understanding of Sarcastic, But Not Sincere, Exchanges

We first tested whether age was related to accurate understanding of video-based social interactions of sarcastic and sincere exchanges. The main effect of age was nonsignificant. We observed a significant main effect of exchange type (Table 1 [*main model*]), such that participants were more accurate for sarcastic than sincere exchanges. Specifically, participants had a higher probability of a correct response for simple sarcasm exchanges (M = 0.93, SE = 0.02) than sincere exchanges (M = 0.85, SE = 0.03), odds ratio (OR) = 0.42, p = .007, 95% CI [0.22, 0.78]; and paradoxical sarcasm exchanges (M = 0.95, SE = 0.02) than sincere exchanges, OR = 0.44, p = .01, 95% CI [0.23, 0.82]. No difference emerged between types of sarcasm, OR = 0.95, p = .89, 95% CI [0.50, 1.81].

The main effect of exchange type was qualified by its interaction with age (Table 1 [*main model*], Figure 1). Older age was unrelated to accuracy for sincere exchanges, b = −0.06, SE = 0.06, z = 0.91, p = .36, 95% CI [−0.18, 0.07]. For both simple and paradoxical sarcasm, however, older age was related to lower accuracy (simple sarcasm: b = −0.60, SE = 0.07, z = 8.12, p < .001, 95% CI [−0.74, −0.45]; paradoxical sarcasm, b = −0.80, SE = 0.07, z = 10.75, p < .001, 95% CI [−0.95, −0.65]). The relationship between age and accuracy was stronger for both types of sarcasm relative to sincerity (simple sarcasm versus sincerity: b = −0.51, SE = 0.06, z = 8.09, p < .001, 95% CI [−0.63, −0.39]; paradoxical sarcasm versus sincerity: b = −0.72, SE = 0.06, z = 11.28, p < .001, 95% CI [−0.85, −0.60]). The relationship between age and accuracy was also stronger for paradoxical sarcasm than simple sarcasm (b = −0.21, SE = 0.07, z = 2.96, p = .003, 95% CI [−0.36, −0.07]). Put another way, younger and middle-aged participants were more accurate for sarcastic versus sincere exchanges, whereas older participants had similar but overall lower accuracy across sincere and sarcastic exchanges (Supplementary Table 2).

The relationship between age and accuracy was stronger for both types of sarcasm relative to sincerity (simple sarcasm versus sincerity: b = −0.51, SE = 0.06, z = 8.09, p < .001, 95% CI [−0.63, −0.39]; paradoxical sarcasm versus sincerity: b = −0.72, SE = 0.06, z = 11.28, p < .001, 95% CI [−0.85, −0.60]). The relationship between age and accuracy was also stronger for paradoxical sarcasm than simple sarcasm (b = −0.21, SE = 0.07, z = 2.96, p = .003, 95% CI [−0.36, −0.07]). Put another way, younger and middle-aged participants were more accurate for sarcastic versus sincere exchanges, whereas older participants had similar but overall lower accuracy across sincere and sarcastic exchanges (Supplementary Table 2).

Better Emotion Recognition Performance Related to More Accurate Understanding of Sarcastic, But Not Sincere, Exchanges

Model results are presented in Table 2 (*emotion recognition model*) and Figure 2A. Emotion recognition performance

Table 1. Probability of a Correct Response: TASIT Part 2 Logistic Linear Mixed Effects Model

Predictors	OR	95% CI	p
(Intercept)	5.52	3.46–8.83	**<0.001**
Age	0.93	0.82–1.06	0.264
Exchange [paradoxical sarcasm]	2.44	1.30–4.60	**0.006**
Exchange [simple sarcasm]	2.65	1.41–5.00	**0.003**
Age × exchange [paradoxical sarcasm]	0.48	0.43–0.55	**<0.001**
Age × exchange [simple sarcasm]	0.60	0.53–0.68	**<0.001**

Notes: CI = confidence interval; OR = odds ratio; TASIT = The Awareness of Social Inference Test. The exchange type reference condition was videos of sincere interactions. Bolded values indicate significant effects (p < .05).

was unrelated to TASIT accuracy overall. Significant two-way interactions, however, emerged between emotion recognition performance and exchange type. Participants with better emotion recognition performance had higher accuracy at detecting sarcasm but not sincerity (simple sarcasm versus sincerity: $b = 0.20$, $SE = 0.06$, $z = 3.21$, $p = .001$, 95% CI [0.08, 0.32]; paradoxical sarcasm versus sincerity: $b = 0.18$, $SE = 0.06$, $z = 2.82$, $p = .005$, 95% CI [−0.05, 0.30]). The relationship between emotion recognition performance and accuracy did not differ between simple and paradoxical sarcasm, $b = 0.02$, $SE = 0.07$, $z = 0.33$, $p = .75$, 95% CI [−0.11, 0.16]. For contrasts of simple slopes versus zero, see Supplementary Table 3.

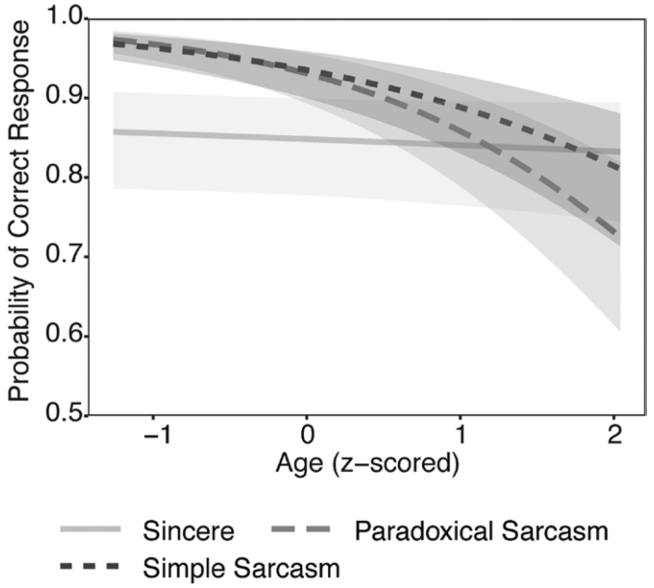

Figure 1. An interaction between age and exchange type emerged when participants engaged in mental state inference from audiovisual stimuli. *Note*: Main model slopes by exchange type. Shading represents the 95% confidence interval. Note the *y*-axis data are plotted on a probability curve for ease of interpretation.

Emotion recognition performance did not, however, moderate the relationship between age and understanding of sarcasm versus sincerity (i.e., no three-way interaction). Given the null result, further model comparison was conducted using Bayes Factor to characterize evidence for a null model (only including main effects and two-way interactions with emotion recognition performance) over the hypothesized model including the three-way interaction. Here, the Bayes Factor reflected decisive evidence (BF01 < .0001; (Held & Ott, 2016) in support of the null model.

Three participants (ages 42, 42, and 65) scored >3 *SD* less than the mean on the RMET, which indicated that they were potential outliers. The two-way interaction reflecting the difference in slopes between sincerity and paradoxical sarcasm became nonsignificant when excluding them, OR = 1.10, $p = .19$, 95% CI [0.95, 1.26]. The difference of the simple slope for paradoxical sarcasm versus zero remained significant; however, $b = 0.19$, $SE = 0.08$, $z = 2.25$, $p = .02$, 95% CI [0.02, 0.35]. The significance and pattern for simple sarcasm and lack thereof for the three-way interaction were unchanged. We therefore conclude that the patterns were relatively unchanged by the inclusion of these participants.

Higher Cognitive Performance Related to More Accurate Understanding of Sarcastic, But Not Sincere, Exchanges

Model results are presented in Table 2 (*cognitive performance model*) and Figure 2B. The logistic linear mixed effects model indicated that cognitive performance was not related to TASIT accuracy overall. Two-way interactions between cognitive performance and exchange type, such that better cognitive performance was related to higher accuracy for sarcastic versus sincere exchanges regardless of age, were qualified by a significant three-way interaction with age.

To decompose the significant three-way interaction, we compared the simple slopes between cognitive performance and exchange-specific accuracy among older (1 *SD* above

Table 2 Probability of a Correct Response: Moderation of the Effect of Age on Accuracy by Emotion Recognition (RMET) and Cognitive (BTACT) Performance Using Logistic Linear Mixed Effects Models

Predictors	Emotion recognition model			Cognitive performance model		
	OR	95% CI	*p*	OR	95% CI	*p*
(Intercept)	5.57	3.48–8.92	**<0.001**	5.75	3.58–9.26	**<0.001**
Age	0.97	0.85–1.11	0.675	0.91	0.80–1.04	0.179
Exchange [paradoxical sarcasm]	2.53	1.34–4.78	**0.004**	2.47	1.30–4.69	**0.006**
Exchange [simple sarcasm]	2.78	1.47–5.26	**0.002**	2.59	1.36–4.92	**0.004**
RMET/BTACT	1.11	0.98–1.27	0.113	0.95	0.77–1.18	0.652
Age × exchange [paradoxical sarcasm]	0.49	0.43–0.56	**<0.001**	0.54	0.48–0.62	**<0.001**
Age × exchange [simple sarcasm]	0.61	0.53–0.70	**<0.001**	0.66	0.58–0.76	**<0.001**
Age × RMET/BTACT	1.05	0.89–1.22	0.570	1.26	0.99–1.60	0.056
Exchange [paradoxical sarcasm] × RMET/BTACT	1.19	1.06–1.35	**0.005**	2.01	1.61–2.50	**<0.001**
Exchange [simple sarcasm] × RMET/BTACT	1.22	1.08–1.38	**0.001**	1.84	1.48–2.28	**<0.001**
Age × exchange [paradoxical sarcasm] × RMET/BTACT	1.02	0.88–1.18	0.788	0.72	0.56–0.92	**0.008**
Age × exchange [simple sarcasm] × RMET/BTACT	1.02	0.88–1.18	0.833	0.65	0.51–0.83	**0.001**

Notes: BTACT = Brief Test of Adult Cognition by Telephone; CI = confidence interval; OR = odds ratio; RMET = Reading the Mind in the Eyes Test. The exchange type reference condition was videos of sincere interactions. Bolded values indicate significant effects ($p < .05$).

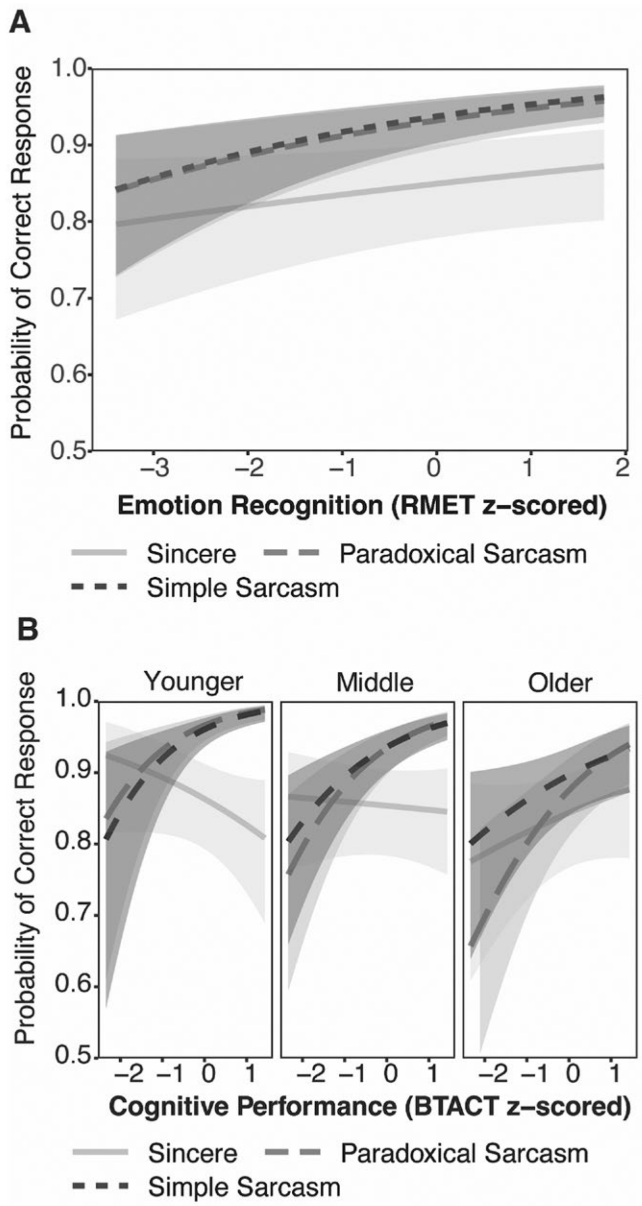

Figure 2. Emotion recognition (**A**) and cognitive (**B**) performance were related to accuracy for sarcastic but not sincere exchanges. BTACT = Brief Test of Adult Cognition by Telephone; RMET = Reading the Mind in the Eyes Test. (**A**) Emotion recognition model slopes by exchange type. Shown as the significant two-way interaction between exchange type and emotion recognition. (**B**) Cognitive performance model slopes by exchange type and age following their significant three-way interaction. Shading represents the 95% confidence interval. Note the *y*-axis data are plotted on a probability curve for ease of interpretation. Significance testing of the simple slopes (vs zero) is presented in Supplementary Table 4.

the sample mean age), middle-aged (sample mean age), and younger participants (1 *SD* below the sample mean age). Contrasts comparing each simple slope to zero are reported in Supplementary Table 5. Among older participants, higher cognitive performance was related to better accuracy for paradoxical sarcasm than sincerity, $b = 0.36$, $SE = 0.13$, $z = 2.89$, $p = .004$, 95% CI [0.11, 0.61]. The slope for simple sarcasm, however, was not different from sincerity, $b = 0.18$, $SE = 0.13$, $z = 1.38$, $p = .17$, 95% CI [−0.08, 0.44]; or paradoxical sarcasm, $b = -0.18$, $SE = 0.14$, $z = 1.35$, $p = .18$, 95% CI [−0.45, 0.08] in these participants. In addition, younger and middle-aged participants showed more differentiation in the relationship between cognitive performance and sarcastic versus sincere exchange accuracy than older participants. Specifically, among younger participants, cognitive performance was more positively related to accuracy for simple sarcasm, $b = 1.03$, $SE = 0.19$, $z = 5.37$, $p < .001$, 95% CI [0.66, 1.41], and paradoxical sarcasm, $b = 1.03$, $SE = 0.20$, $z = 5.12$, $p < .001$, 95% CI [0.64, 1.42], than sincerity. The slopes for paradoxical sarcasm and simple sarcasm were not different in younger participants, $b = -0.004$, $SE = 0.23$, $z = 0.02$, $p = .99$, 95% CI [−0.45, 0.45]. The same pattern emerged for middle-aged participants, such that cognitive performance was more strongly positively related to accuracy for simple, $b = 0.61$, $SE = 0.11$, $z = 5.53$, $p < .001$, 95% CI [0.39, 0.82], and paradoxical, $b = 0.70$, $SE = 0.11$, $z = 6.22$, $p < .001$, 95% CI [0.48, 0.92], sarcasm than sincerity. Among middle-aged participants, the simple slope for paradoxical sarcasm was also not different from simple sarcasm, $b = 0.09$, $SE = 0.12$, $z = 0.72$, $p = .47$, 95% CI [−0.15, 0.33]. For contrasts of simple slopes versus zero, see Supplementary Table 4. An exploratory depiction of the relationship for each BTACT subtest is shown in Supplementary Figure 1.

Delineating Contributions of Emotion Recognition and Cognitive Performance for More Accurate Understanding of Sarcastic Versus Sincere Exchanges

It is possible that the association between emotion recognition and sarcasm understanding was driven by a shared association with good cognitive ability (Supplementary Table 1). To test this possibility, we conducted logistic linear mixed effects models as noted previously, including both RMET and BTACT scores in the same model. A model allowing interactions between RMET and BTACT did not improve model fit measured by AIC as compared with a model where RMET and BTACT interacted with age and exchange type but with not each other, $X2$ (1, $N = 238$) = 8.01, $p = .24$ (Supplementary Table 5). Relevant here, the significant two-way interactions between cognitive performance and exchange type as well as significant three-way interactions between cognitive performance, exchange type, and age retained significance in the previously reported patterns. For emotion recognition, the two-way interactions between RMET and exchange type were not significant in the model including the BTACT, unlike before. The three-way interactions between the RMET, exchange type, and age were nonsignificant as before.

Discussion

Our findings show that accuracy at understanding sarcasm, but not sincerity, was negatively related to older age across our adult lifespan sample. As in prior investigations using the TASIT (McDonald et al., 2018; Phillips et al., 2015), sincere exchanges were more difficult compared with sarcastic exchanges for all ages. Also, regardless of age, both facial emotion recognition and cognitive performance related to a more accurate understanding of sarcastic, but not sincere, exchanges. Age did not further moderate the interaction between facial emotion recognition and exchange type. We did observe, however, that the differential association of cognitive performance and accuracy by exchange type was weaker among older versus younger and middle-aged adults.

Together, these findings support that both emotion recognition from faces and cognitive performance facilitate complex mental state inference and suggest a more nuanced account for whether age deficits in these abilities underlie age deficits in sarcasm understanding as opposed to mental state understanding more broadly.

Although the TASIT was originally developed to measure social dysfunction in the context of traumatic brain injury, previous investigations have used it to study mental state inference over the adult lifespan (Lavrencic et al., 2016; McDonald et al., 2018), between age cohorts (Burdon et al., 2016; Grainger et al., 2018; Phillips et al., 2015), and between older age-matched clinical and nonclinical cohorts (e.g., Alzheimer's disease and other dementias [Bora et al., 2015; Kumfor et al., 2017; Rankin et al., 2009; Shany-Ur et al., 2012]; Huntington's disease [Larsen et al., 2016; Parkinson's disease: Pell et al., 2014]). Given the heterogeneity of scoring methods (e.g., across or between exchange types) and populations, a strength of the current work was characterizing exchange-specific differences in one of the largest adult lifespan samples to date. By doing so, we support that the TASIT is sensitive to age differences in mental state understanding, specifically under the context of incongruity between literal versus nonliteral interpretation. Albeit the older age participants in our sample were still quite good at the task. Studies that collapse across exchange types may thus be less sensitive to age and cohort differences than those that study sarcasm specifically. To this point, it may be puzzling to reconcile that participants were worse at understanding sincere compared with sarcastic exchanges with the premise that sarcasm is a more complex mental state. In the TASIT, sarcastic exchanges were designed to be "explicit and unambiguous" (McDonald et al., 2018), which may not always be true in other measures or in real-world social interactions. Complexity, as defined by requiring individuals to interpret and integrate additional cues for accurate inference, did not make sarcastic exchanges in the TASIT more difficult per se. Rather, those who could not engage with this complexity successfully (e.g., those with poorer cognitive performance) had more difficulty with sarcastic exchanges. Alternately, the nature of the task may have made participants more suspicious to nonliteral meaning in the sincere exchanges, thereby reducing overall performance relative to sarcastic exchanges. Disentangling sarcasm types, as in the current work, may provide more refined insight about mental state complexity and contributing individual differences.

Specifically, we demonstrate some of the first evidence among an adult lifespan sample of a more pronounced age-related decline in understanding of paradoxical versus simple sarcasm, which differs from previous reports' finding of no difference between sarcasm types (McDonald et al., 2003; Phillips et al., 2015). Simple sarcasm was originally conceived as more complex because the literal meaning of the words is still plausible (albeit contrasting nonverbal cues), as compared with paradoxical sarcasm (relying more on semantic processing; [McDonald et al., 2003]). A prior report in young, middle-aged, and older adults found across age groups that paradoxical sarcasm was easier to understand than simple sarcasm, an effect that was not observed in our data. These conflicting findings across studies point to the need to better characterize if and how simple sarcasm is more complex or difficult than paradoxical sarcasm. Regarding sincere exchanges, McDonald et al. (2018) found the opposite pattern; a decrease from adulthood to older adulthood was more pronounced for sincere than sarcastic exchanges. Future work reconciling these age trends is necessary, although we again conclude that exchange type matters for the interpretation of age and/or cohort differences.

In contrast to prior work (Phillips et al., 2015), we did not find evidence that difficulties in emotion recognition moderated the relationship between age and sarcasm understanding from audiovisual social interactions. An important distinction that we made here was to examine the relative contribution of emotion recognition performance to understanding of sarcastic versus sincere exchanges. Emotion recognition did not appear to facilitate accurate mental state inferences overall or for older versus younger adults. Rather, better emotion recognition facilitated better understanding of sarcasm but not sincerity—despite that sincerity was overall more difficult to interpret. This is a key finding because it suggests that the ability to understand emotions from faces may become especially useful in complex scenarios, such as when the literal interpretation of the speaker's words is not what they mean. This interpretation is tempered because emotion recognition did not have an independent influence on sarcasm understanding when accounting for cognitive performance. As such, good cognitive performance may be needed to take advantage of the multiplicity of cues from videos of social interactions.

A limitation of the current work is that we assessed facial emotion recognition using the RMET, which is widely used but criticized for its limitations regarding psychometric validity and generalizability (Higgins et al., 2023; Kim et al., 2024). Although the RMET has sensitivity to variation in healthy aging (Henry et al., 2013), it may not accurately capture the facial emotion recognition that participants engage in during the TASIT. The strength and direction of the association between emotion recognition and sarcasm understanding reported here is similar to other adult lifespan or aging samples using dynamic stimuli (Lavrencic et al., 2015; Phillips et al., 2015) but may have emerged for different reasons. A dynamic facial emotion recognition task (Gonçalves et al., 2018; Hayes et al., 2020; Holland et al., 2019; Phillips et al., 2015) may have been more sensitive to age-specific contributions of emotion recognition performance in this context. Another possible interpretation is that, given that multiple modalities can inform emotion recognition, difficulty in recognizing emotions from faces could be overcome by older adults' preserved competence in other modalities that we did not test (e.g., vocal emotion recognition). Supporting this idea, adolescents' emotion recognition from voices follows a different developmental trajectory than emotion recognition from faces (Morningstar et al., 2018), and the same may be possible for protracted declines in older adulthood (Ruffman et al., 2008). Even so, difficulties in emotion recognition may not limit the performance of participants to use or integrate other cues to make accurate mental state inferences.

Cognitive processes facilitate socioemotional cue integration (Fiske & Taylor, 1991; Hamilton et al., 2022; Lighthall, 2020; Moran, 2013; Sandoz et al., 2014). This finding may explain why, in the current work, better cognitive performance was related to accurate mental state inference about sarcastic but not sincere exchanges. For example, during sarcastic exchanges, participants with good cognitive ability may have more successfully inhibited the distracting literal interpretation of the speaker's words to instead use different facial, paralinguistic, and contextual cues to determine the speaker's intended meaning. Or, retrieving information from different cues to respond to four subsequent questions probing the speaker's thoughts, feelings, and intentions could have required good working memory for holding in mind or retrieving speakers' words, expressions, and context (Krendl et al., 2022). That we did not observe a differential contribution of cognitive performance by type of sarcasm suggests that each type may be similarly cognitively demanding but for different reasons. Supplemental exploratory analysis revealed a consistent pattern of effects across the executive functioning and memory subtests of the cognitive performance battery (Lachman et al., 2014) used in this study, although the three-way interactions with age were more heterogenous.

We predicted that older adults would rely more than younger or middle-aged adults on their cognitive performance to accurately decode sarcasm versus sincerity. Unexpectedly, we found that older adults similarly relied on cognitive performance during more (simple sarcasm) and less (sincerity) complex exchanges. Whether mental state understanding is a specialized form of cognition is debated (Moran, 2013). Our results may indicate that specialization may differ across the adult lifespan, consistent with findings that age-related declines in mental state understanding are not fully independent from age-related declines in cognition (Moran, 2013). These findings may also be due to ceiling effects at sarcasm understanding among the younger adults, which should be noted as a potential limitation for using the TASIT in cognitively healthy younger and middle-aged individuals. Also, the TASIT videos were developed approximately two decades ago, and the actor demographics are unreported (McDonald et al., 2003). Extant work suggests that perceiver and target demographics (e.g., race, age) interact to influence accuracy at understanding others' mental states (Cassidy et al., 2021; Kim et al., 2024). Thus, the previous observation of no overall effect of participant sex/gender (McDonald et al., 2018), for instance, does not rule out more subtle patterns that may emerge when considering not only who is the perceiver but also who is the speaker and whether the speaker is being sarcastic or not. Future work should therefore determine whether demographic differences between perceivers and targets could explain age-related difficulties in sarcasm understanding. These potential limitations demonstrate the necessity of new or updated naturalistic paradigms for complex mental state understanding that are inclusive and valid across a range of participant and target demographics and avoid ceiling effects.

In summary, our results showed that the potential benefits of dynamic stimuli for reducing or eliminating age differences in mental state understanding (Hayes et al., 2020; Holland et al., 2019) may be tempered by the added complexity of naturalistic paradigms that incorporate multiple types of mental state cues. Moreover, this result was specifically associated with complex mental state inference for sarcasm (vs sincerity), where listeners could not solely rely on the literal meaning of the speaker's words. We also extended observations in prior work (Phillips et al., 2015) by showing that general cognitive performance, as opposed to working memory specifically, was related to better understanding of sarcasm (vs sincerity). These findings motivate important theoretical questions for future research on mental state inference among older adults. Do multiple modality stimuli allow older compared with younger adults to use different cues to facilitate performance? Or, does the gestalt of cues (i.e., how they are cognitively integrated) promote accurate mental state inference even if older adults' worse sociocognitive abilities degrade their understanding of individual cues? Naturalistic paradigms that allow for individuals to use more varied strategies may better capture older adults' cognitive and behavioral heterogeneity (Fittipaldi et al., 2024; Lowsky et al., 2014). From a practical standpoint, we can learn not only which older adults are successful in social interactions but also why. That is, we can learn if they are using similar or different strategies as younger adults (Krendl et al., 2022). Older adults with good cognitive function may benefit from training similar strategies as younger adults (e.g., [Lecce et al., 2019]), whereas older adults with impaired cognitive function may benefit from interventions that do not rely on good cognitive function (e.g., [Krendl & Hughes, 2024]). By personalizing the intervention aims to individuals, we can better address limitations to having positive social interactions, such as misunderstanding others' thoughts and feelings. Better understanding of others relates to feelings of closeness and relationship quality (Lecce et al., 2017) that may, in turn, serve the longer-term goal of reducing loneliness and its pernicious consequences in aging (Cacioppo et al., 2014).

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

Funding

This work was supported by the National Institute on Aging of the National Institutes of Health (1R01AG057764 to N. C. Ebner and R. N. Spreng, R01AG068563 to R. N. Spreng, and R01AG072658 to N. C. Ebner); the Florida Department of Health Ed and Ethel Moore Alzheimer's Disease Research Program grant (22A10 to N. C. Ebner); Canadian Institutes of Health Research postdoctoral fellowship (181831 to C. S. Hughes); Fonds de Recherche du Quebec postdoctoral fellowship (#320680 to C. S. Hughes); and Scientific Research Network on Decision Neuroscience & Aging collaboration award (to C. S. Hughes and N. C. Ebner).

Conflict of Interest

None.

Data Availability

De-identified data and analytic code needed to reproduce analyses are provided at https://osf.io/td3eg/. Materials are cited and described in the text. The Awareness of Social Inference Test (TASIT) can be retrieved from https://www.assbi.com.au/The-Awareness-of-Social-Inference-Test-(TASIT). Given the secondary data-analytic nature of this research, the study design, hypotheses, and analytic plan were not preregistered.

References

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test Revised Version: A study with normal adults, and adults with Asperger Syndrome or High-functioning Autism. *Journal of Child Psychology and Psychiatry*, 42(2), 241–251. https://doi.org/10.1017/s0021963001006643

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Bora, E., Walterfang, M., & Velakoulis, D. (2015). Theory of mind in behavioural-variant frontotemporal dementia and Alzheimer's disease: A meta-analysis. *Journal of Neurology, Neurosurgery, and Psychiatry*, 86(7), 714–719. https://doi.org/10.1136/jnnp-2014-309445

Burdon, P., Dipper, L., & Cocks, N. (2016). Exploration of older and younger British adults' performance on The Awareness of Social Inference Test (TASIT). *International Journal of Language & Communication Disorders*, 51(5), 589–593. https://doi.org/10.1111/1460-6984.12233

Cacioppo, S., Capitanio, J. P., & Cacioppo, J. T. (2014). Toward a neurology of loneliness. *Psychological Bulletin*, 140(6), 1464–1504. https://doi.org/10.1037/a0037618

Cassidy, B. S., Hughes, C., & Krendl, A. C. (2021). Age differences in neural activity related to mentalizing during person perception. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 28(1), 143–160. https://doi.org/10.1080/13825585.2020.1718060

Charlton, R. A., Barrick, T. R., Markus, H. S., & Morris, R. G. (2009). Theory of mind associations with other cognitive functions and brain imaging in normal aging. *Psychology and Aging*, 24(2), 338–348. https://doi.org/10.1037/a0015225

Cornwell, E. Y., & Waite, L. J. (2009). Social disconnectedness, perceived isolation, and health among older adults. *Journal of Health and Social Behavior*, 50(1), 31–48. https://doi.org/10.1177/002214650905000103

Ebner, N. C., & Johnson, M. K. (2009). Young and older emotional faces: Are there age group differences in expression identification and memory? *Emotion*, 9(3), 329–339. https://doi.org/10.1037/a0015179

Ekman, P. (2006). *Darwin and facial expression: A century of research in review*. ISHK.

English, T., & Carstensen, L. L. (2014). Selective narrowing of social networks across adulthood is associated with improved emotional experience in daily life. *International Journal of Behavioral Development*, 38(2), 195–202. https://doi.org/10.1177/0165025413515404

Fernandes, C., Barbosa, F., Martins, I. P., & Marques-Teixeira, J. (2021). Aging and social cognition: A comprehensive review of the literature. *Psychology & Neuroscience*, 14(1), 1–15. https://doi.org/10.1037/pne0000251

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: Psychological Sciences and Social Sciences*, 72(1), 71–81. https://doi.org/10.1093/geronb/gbw088

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed., pp. xviii, 717). Mcgraw-Hill Book Company.

Fittipaldi, S., Legaz, A., Maito, M., Hernandez, H., Altschuler, F., Canziani, V., Moguilner, S., Gillan, C. M., Castillo, J., Lillo, P., Custodio, N., Avila-Funes, J. A., Cardona, J. F., Slachevsky, A., Henriquez, F., Fraile-Vazquez, M., Cruz de Souza, L., Borroni, B., Hornberger, M., … Ibanez, A. (2024). Heterogeneous factors influence social cognition across diverse settings in brain health and age-related diseases. *Nature Mental Health*, 2(1), 63–75. https://doi.org/10.1038/s44220-023-00164-3

Gonçalves, A. R., Fernandes, C., Pasion, R., Ferreira-Santos, F., Barbosa, F., & Marques-Teixeira, J. (2018). Effects of age on the identification of emotions in facial expressions: A meta-analysis. *PeerJ*, 6, e5278. https://doi.org/10.7717/peerj.5278

Grainger, S. A., Henry, J. D., Steinvik, H. R., Vanman, E. J., Rendell, P. G., & Labuschagne, I. (2018). Intranasal oxytocin does not reduce age-related difficulties in social cognition. *Hormones and Behavior*, 99, 25–34. https://doi.org/10.1016/j.yhbeh.2018.01.009

Gurera, J. W., & Isaacowitz, D. M. (2019). Chapter 14 – Emotion regulation and emotion perception in aging: A perspective on age-related differences and similarities. In N. Srinivasan (Ed.), *Progress in brain research* (Vol. 247, pp. 329–351). Elsevier. https://doi.org/10.1016/bs.pbr.2019.02.007

Hamilton, L. J., Gourley, A. N., & Krendl, A. C. (2022). They cannot, they will not, or we are asking the wrong questions: Re-examining age-related decline in social cognition. *Frontiers in Psychology*, 13, 894522. https://doi.org/10.3389/fpsyg.2022.894522

Hayes, G. S., McLennan, S. N., Henry, J. D., Phillips, L. H., Terrett, G., Rendell, P. G., Pelly, R. M., & Labuschagne, I. (2020). Task characteristics influence facial emotion recognition age-effects: A meta-analytic review. *Psychology and Aging*, 35(2), 295–315. https://doi.org/10.1037/pag0000441

Held, L., & Ott, M. (2016). How the maximal evidence of p-values against point null hypotheses depends on sample size. *The American Statistician*, 70(4), 335–341. https://doi.org/10.1080/00031305.2016.1209128

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), 826–839. https://doi.org/10.1037/a0030677

Hess, T. M. (2014). Selective engagement of cognitive resources: Motivational influences on older adults' cognitive functioning. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 9(4), 388–407. https://doi.org/10.1177/1745691614527465

Higgins, W. C., Kaplan, D. M., Deschrijver, E., & Ross, R. M. (2023). Construct validity evidence reporting practices for the Reading the mind in the eyes test: A systematic scoping review. *Clinical Psychology Review*, 108, 102378. https://doi.org/10.1016/j.cpr.2023.102378

Holland, C. A. C., Ebner, N. C., Lin, T., & Samanez-Larkin, G. R. (2019). Emotion identification across adulthood using the Dynamic FACES database of emotional expressions in younger, middle aged, and older adults. *Cognition and Emotion*, 33(2), 245–257. https://doi.org/10.1080/02699931.2018.1445981

Huang, L., & Galinsky, A. D. (2023). SHARP wit: Why receiving sarcasm improves perspective-taking. *Current Opinion in Psychology*, 54, 101709. https://doi.org/10.1016/j.copsyc.2023.101709

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. https://doi.org/10.1016/j.neuroimage.2019.02.024

Kim, H. A., Kaduthodil, J., Strong, R. W., Germine, L. T., Cohan, S., & Wilmer, J. B. (2024). Multiracial Reading the Mind in the Eyes Test (MRMET): An inclusive version of an influential measure. *Behavior Research Methods*, 56, 5900–5917. https://doi.org/10.3758/s13428-023-02323-x

Krendl, A. C., & Hughes, C. S. (2024). Mindsets over matter: priming theory of mind improves older adults' mental state attributions about naturalistic social interactions. *Aging & Mental Health*, 1–9. https://doi.org/10.1080/13607863.2024.2399091

Krendl, A. C., Mannering, W., Jones, M. N., Hugenberg, K., & Kennedy, D. P. (2022). Determining whether older adults use similar strategies to young adults in theory of mind tasks. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 78(6), 969–976. https://doi.org/10.1093/geronb/gbac187

Kumfor, F., Honan, C., McDonald, S., Hazelton, J. L., Hodges, J. R., & Piguet, O. (2017). Assessing the "social brain" in dementia: Applying TASIT-S. *Cortex*, 93, 166–177. https://doi.org/10.1016/j.cortex.2017.05.022

Lachman, M. E., Agrigoroaei, S., Tun, P. A., & Weaver, S. L. (2014). Monitoring cognitive functioning: Psychometric properties of the brief test of adult cognition by telephone. *Assessment*, 21(4), 404–417. https://doi.org/10.1177/1073191113508807

Larsen, I. U., Vinther-Jensen, T., Gade, A., Nielsen, J. E., & Vogel, A. (2016). Do I misconstrue? Sarcasm detection, emotion recognition, and theory of mind in Huntington disease. *Neuropsychology*, 30(2), 181–189. https://doi.org/10.1037/neu0000224

Lavrencic, L. M., Kurylowicz, L., Valenzuela, M. J., Churches, O. F., & Keage, H. A. D. (2016). Social cognition is not associated with cognitive reserve in older adults. *Aging, Neuropsychology and Cognition*, 23(1), 61–77. https://doi.org/10.1080/13825585.2015.1048773

Lecce, S., Ceccato, I., Bianco, F., Rosi, A., Bottiroli, S., & Cavallini, E. (2017). Theory of Mind and social relationships in older adults: The role of social motivation. *Aging & Mental Health*, 21(3), 253–258. https://doi.org/10.1080/13607863.2015.1114586

Lecce, S., Ceccato, I., & Cavallini, E. (2019). Investigating ToM in aging with the MASC: From accuracy to error type. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 26(4), 541–557. https://doi.org/10.1080/13825585.2018.1500996

Lighthall, N. R. (2020). Neural mechanisms of decision-making in aging. *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(1), e1519. https://doi.org/10.1002/wcs.1519

Lowsky, D. J., Olshansky, S. J., Bhattacharya, J., & Goldman, D. P. (2014). Heterogeneity in healthy aging. *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, 69(6), 640–649. https://doi.org/10.1093/gerona/glt162

McDonald, S. (2012). New frontiers in neuropsychological assessment: Assessing social perception using a standardised instrument, the awareness of social inference test. *Australian Psychologist*, 47(1), 39–48. https://doi.org/10.1111/j.1742-9544.2011.00054.x

McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., & Neulinger, K. (2006). Reliability and validity of The Awareness of Social Inference Test (TASIT): A clinical test of social perception. *Disability and Rehabilitation*, 28(24), 1529–1542. https://doi.org/10.1080/09638280600646185

McDonald, S., Flanagan, S., & Rollins, J. (n.d.). *The Awareness of Social Inference Test (TASIT)*. https://www.assbi.com.au/The-Awareness-of-Social-Inference-Test-(TASIT)

McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). A new clinical tool for assessing social perception after traumatic brain injury. *Journal of Head Trauma Rehabilitation*, 18(3), 219–238. https://doi.org/10.1097/00001199-200305000-00001

McDonald, S., Honan, C., Allen, S. K., El-Helou, R., Kelly, M., Kumfor, F., Piguet, O., Hazelton, J. L., Padgett, C., & Keage, H. A. D. (2018). Normal adult and adolescent performance on TASIT-S, a short version of The Assessment of Social Inference Test. *The Clinical Neuropsychologist*, 32(4), 700–719. https://doi.org/10.1080/13854046.2017.1400106

Moran, J. M. (2013). Lifespan development: The effects of typical aging on theory of mind. *Behavioural Brain Research*, 237(1), 32–40. https://doi.org/10.1016/j.bbr.2012.09.020

Morningstar, M., Nelson, E. E., & Dirks, M. A. (2018). Maturation of vocal emotion recognition: Insights from the developmental and neuroimaging literature. *Neuroscience and Biobehavioral Reviews*, 90, 221–230. https://doi.org/10.1016/j.neubiorev.2018.04.019

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, 60, 173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656

Pell, M. D., Monetta, L., Rothermich, K., Kotz, S. A., Cheang, H. S., & McDonald, S. (2014). Social perception in adults with Parkinson's disease. *Neuropsychology*, 28(6), 905–916. https://doi.org/10.1037/neu0000090

Phillips, L. H., Allen, R., Bull, R., Hering, A., Kliegel, M., Phillips, L. H., Allen, R., Bull, R., Hering, A., Kliegel, M., & Channon, S. (2015). Older adults have difficulty in decoding sarcasm older adults have difficulty in decoding sarcasm. *Developmental Psychology*, 51(12), 1840–1852. https://doi.org/10.1037/dev0000063

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

Polk, R., Horta, M., Lin, T., Porges, E., Ojeda, M., Nazarloo, H. P., Carter, C. S., & Ebner, N. C. (2022). Evaluating the neuropeptide–social cognition link in ageing: The mediating role of basic cognitive skills. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1858), 20210048. https://doi.org/10.1098/rstb.2021.0048

Rankin, K. P., Salazar, A., Gorno-Tempini, M. L., Sollberger, M., Wilson, S. M., Pavlic, D., Stanley, C. M., Glenn, S., Weiner, M. W., & Miller, B. L. (2009). Detecting sarcasm from paralinguistic cues: Anatomic and cognitive correlates in neurodegenerative disease. *Neuroimage*, 47(4), 2005–2015. https://doi.org/10.1016/j.neuroimage.2009.05.077

Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information. *Psychology and Aging*, 29(1), 1–15. https://doi.org/10.1037/a0035194

Riediger, M., Voelkle, M. C., Ebner, N. C., & Lindenberger, U. (2011). Beyond "happy, angry, or sad?": Age-of-poser and age-of-rater effects on multi-dimensional emotion perception. *Cognition and Emotion*, 25(6), 968–982. https://doi.org/10.1080/02699931.2010.540812

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews*, 32(4), 863–881. https://doi.org/10.1016/j.neubiorev.2008.01.001

Sandoz, M., Démonet, J.-F., & Fossard, M. (2014). Theory of mind and cognitive processes in aging and Alzheimer type dementia: A systematic review. *Aging & Mental Health*, 18(7), 815–827. https://doi.org/10.1080/13607863.2014.899974

Shamay-Tsoory, S. G., Tomer, R., & Aharon-Peretz, J. (2005). The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology*, 19(3), 288–300. https://doi.org/10.1037/0894-4105.19.3.288

Shany-Ur, T., Poorzand, P., Grossman, S. N., Growdon, M. E., Jang, J. Y., Ketelle, R. S., Miller, B. L., & Rankin, K. P. (2012). Comprehension of insincere communication in neurodegenerative disease: Lies, sarcasm, and theory of mind. *Cortex*, 48(10), 1329–1341. https://doi.org/10.1016/j.cortex.2011.08.003

Spreng, R. N., & Turner, G. R. (2019). The shifting architecture of cognition and brain function in older adulthood. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 14(4), 523–542. https://doi.org/10.1177/1745691619827511

Wang, Z., & Su, Y. (2013). Age-related differences in the performance of theory of mind in older adults: A dissociation of cognitive and affective components. *Psychology and Aging*, 28(1), 284–291. https://doi.org/10.1037/a0030876

Winner, E., & Leekam, S. (1991). Distinguishing irony from deception: Understanding the speaker's second-order intention. *British Journal of Developmental Psychology*, 9(2), 257–270. https://doi.org/10.1111/j.2044-835x.1991.tb00875.x

Zaitchik, D. (1990). When representations conflict with reality: The preschooler's problem with false beliefs and 'false' photographs. *Cognition*, 35(1), 41–68. https://doi.org/10.1016/0010-0277(90)90036-j