Mindreading From the Eyes Declines With Aging – Evidence From 1,603 Subjects

Jana Kynast[1,2], Eva Maria Quinque[1,2], Maryna Polyakova[1,2], Tobias Luck[3], Steffi G. Riedel-Heller[2,4], Simon Baron-Cohen[5], Andreas Hinz[6], A. Veronica Witte[1,2], Julia Sacher[1,2,7], Arno Villringer[1,2,7] and Matthias L. Schroeter[1,2,7]*

[1] Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [2] LIFE – Leipzig Research Center for Civilization Diseases, University of Leipzig, Leipzig, Germany, [3] Faculty of Applied Social Sciences, University of Applied Sciences Erfurt, Erfurt, Germany, [4] Institute for Social Medicine, Occupational Medicine and Public Health, University Hospital Leipzig, Leipzig, Germany, [5] Department of Psychiatry, Autism Research Centre, University of Cambridge, Cambridge, United Kingdom, [6] Department for Medical Psychology and Sociology, University Hospital Leipzig, Leipzig, Germany, [7] Clinic for Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany

Social cognition, in particular mindreading, enables the understanding of another individual's feelings, intentions, desires, and mental states. The Reading the Mind in the Eyes Test (RMET) captures the ability to identify mental states from gaze. We investigated RMET accuracy in the context of age and cognition across the whole adult age-range (19–79 years) in a very large population-based sample (N = 1,603) with linear regression models accounting for cognitive abilities, neurological diseases, and psychiatric disorders. Higher age predicted lower RMET performance in women and men, suggesting difficulties to infer mental states from gaze at older age. Effects remained stable when taking other cognitive abilities and psychiatric disorders or neurological diseases into account. Our results show that RMET performance as a measure of social cognition declines with increasing age.

Keywords: aging, Reading the Mind in the Eyes Test, social cognition, women, men

INTRODUCTION

Theory of Mind (ToM) is an essential socio-cognitive ability that enables the attribution of mental states to self and others (Premack and Woodruff, 1978). The knowledge about another individual's intentions, desires, and beliefs facilitates the prediction of future behavior, and enables a socially adequate reaction. Thus, ToM, also referred to as "mentalizing" (Morton and Johnson, 1991) or "mindreading" (Whiten, 1991), is essential for successful social interactions. On the contrary, deficits in ToM can compromise the initiation and maintenance of social relationships and individual well-being. ToM impairment can be found in several mental conditions, such as autism

---

Abbreviations: LIFE, Leipzig research Centre for Civilization Diseases; RMET, Reading the Mind in the Eyes test; ToM, Theory of Mind; WST, Wortschatztest.

(Baron-Cohen et al., 1985, 1997, 2015; Brune and Brune-Cohrs, 2006), but also neurodegenerative diseases, like behavioral variant frontotemporal dementia (Gregory et al., 2002; Schroeter, 2012; Pardini et al., 2013; Schroeter et al., 2014, 2018), Alzheimer's disease, and mild cognitive impairment (Baglio et al., 2012; Poletti and Bonuccelli, 2013). Recently, the Diagnostic and Statistical Manual of Mental Diseases, 5th edition (DSM-5; American Psychiatric Association, 2013) introduced social cognition as one core domain for the evaluation of cognitive performance in the context of dementia and its pre-stages, i.e., mild and major neurocognitive disorder, highlighting the clinical relevance of the reliable and effective assessment of ToM in adults.

The "Reading the Mind in the Eyes" Test (RMET; Baron-Cohen et al., 1997, 2001) is an established instrument for the neuropsychological assessment of socio-cognitive functions. It captures the ability to attribute mental states from gaze. The revised version includes 36 photographs of the eye-region of a person (Baron-Cohen et al., 2001). Out of four response options, the term most appropriately describing the pictured mental state shall be selected. Translated into many languages, the RMET is one of the most frequently used tasks to investigate ToM in adults (Kirkland et al., 2013). As the typical accuracy rate is usually below 100% (Pardini and Nichelli, 2009), the RMET is not limited by ceiling effects. The RMET has been discussed to tap mainly "hot" or "affective" ToM, which requires an understanding of others' emotions, affective states, or feelings, whereas "cold" or "cognitive" ToM requires an understanding of their cognitive states, beliefs, thoughts, or intentions (Henry et al., 2013).

Recently, the influence of age on RMET accuracy has been in the focus of interest. Studies investigating age-related differences using cross-sectional data have shown a reduced RMET accuracy at older age (Pardini and Nichelli, 2009; Baglio et al., 2012; El Haj et al., 2016; Warrier et al., 2018). Other tests of ToM and social cognition show similar age effects. Henry et al. (2013) conducted a meta-analysis across six different types of ToM tasks, i.e., stories, eyes, videos, false belief-video, false belief-other, and faux pas, categorized into "affective" and "cognitive" ToM. Overall, older adults performed more poorly than younger adults with moderate magnitude. Remarkably, no effects for specific task parameters were detected.

Our study aimed at investigating the impact of age and several cognitive functions on the ability to identify mental states from gaze in the RMET in a large, population-based sample comprising 1,603 subjects aged 19–79 years. Based on literature findings (see above), we predicted better RMET performance in younger than older adults. Regarding cognitive abilities, RMET accuracy was hypothesized to be positively related to verbal intelligence (measured with a vocabulary test, and estimating individual vocabulary abilities) as suggested recently (Ahmed and Miller, 2011; Peterson and Miller, 2012; Baker et al., 2014). We further tested the impact of attention, memory, executive functions, and verbal fluency on RMET performance across the study sample, because age was reported to interact with these cognitive domains and RMET performance (Bailey and Henry, 2008; Duval et al., 2011; Dal Monte et al., 2014). Yet, the effect of domain-specific cognitive abilities on RMET accuracy has never been investigated

in a large, community-dwelling sample covering the complete adult age range.
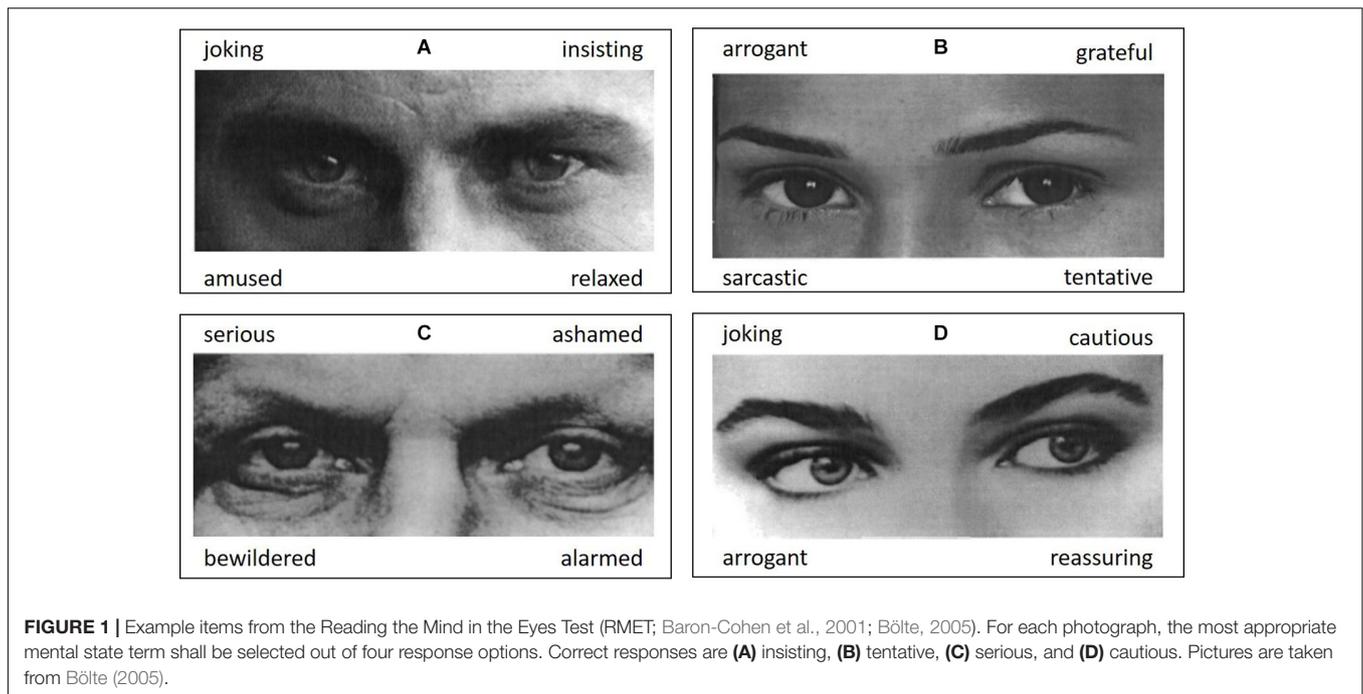
MATERIALS AND METHODS

Psychological Tests

This study was part of the adult study of the Leipzig Research Centre for Civilization Diseases (LIFE). The detailed study procedure is described elsewhere (Loeffler et al., 2015). ToM was assessed with the German version of the revised RMET (Baron-Cohen et al., 2001; Bölte, 2005). The test contains 36 pictures showing the eyes of either a man or a woman. Every picture is presented with four adjectives of which the one describing the pictured mental state best should be selected. **Figure 1** illustrates four test stimuli. Computerized assessment was self-paced and needed approximately 10–15 min for completion. Additionally, all participants completed a 42-item test of verbal intelligence (vocabulary test, so called Wortschatztest, WST; Schmidt and Metzler, 1992). For each item, six words were presented. Of those, five were pseudo words and only one a real term existing in German. The latter shall be selected. For both tests (RMET, WST), accuracy rates (percentage of correct responses) were analyzed. Reaction times were not analyzed, as subjects were not explicitly instructed to respond as fast as possible. Furthermore, we avoided the well-known bias of speed slowing due to aging (Schroeter et al., 2003) by abstaining from analyses of reaction times.

Study Cohort

All study participants were randomly selected from the population according to the study protocol (Loeffler et al., 2015). Of those randomly sampled citizens, where data had been provided by the resident's registration office of the city of Leipzig, 10,000 participants were investigated in total. A subsample was characterized by in-depth cognitive phenotyping including, beside several other tests, the RMET. As the LIFE study was localized in the city of Leipzig (Saxony/Germany), most participants represented Caucasians from diverse social backgrounds. Information on age, and psychological measures (description below) was assessed.

Information on affiliation to either women or men was obtained directly from the resident's registration office of the city of Leipzig, and, additionally, with a questionnaire given to the participants containing the question: "Which gender are you belonging to?" (In German "Welches Geschlecht haben Sie?"). The final study sample included 1,603 persons aged between 19 and 79 years (905 men; **Table 1**). Of note, both ways of determining affiliation to women or men showed a high coincidence, i.e., 100% of the study population with available information from both sources (93.4%) did show agreement on both factors. For the total LIFE population of 10,000 participants, this was the case for 99.93% (90% with both information). None of the six participants with disagreement between both measures in the whole LIFE study were included in our study. In the following, this factor – affiliation to women or men – will be characterized by the term gender. Please, find a critical discussion of this approach in the discussion section.

FIGURE 1 | Example items from the Reading the Mind in the Eyes Test (RMET; Baron-Cohen et al., 2001; Bölte, 2005). For each photograph, the most appropriate mental state term shall be selected out of four response options. Correct responses are **(A)** insisting, **(B)** tentative, **(C)** serious, and **(D)** cautious. Pictures are taken from Bölte (2005).

All individuals scored above chance level in either psychological test (WST > 17%; RMET > 25%), ensuring a sufficient understanding of the test instructions. Chance levels for each test have been calculated by relating the number of correct responses to the total number of responses provided for each item. Since in the RMET each item is presented with four response alternatives of which only one answer is correct, chance level is calculated as 1/4 (25%, respectively). In the WST, for each item six response alternatives are given of which only one is correct. The chance level is hence calculated as 1/6 (17%, respectively).

The study was approved by the ethics committee of the University of Leipzig and was in accordance with the latest version of the Declaration of Helsinki. Each subject provided written informed consent.

Statistical Analysis
Linear Regression of Age and Gender on RMET Performance
The influence of age, gender and verbal intelligence on RMET performance was tested with a linear regression model predicting RMET accuracy (criterion) from age, gender, the interaction term age*gender, and WST performance (predictors). We predicted older age to be linearly associated with lower RMET performance. The interaction term (age*gender) tests whether age affects RMET performance differently in men and women. Considering findings from the literature, a small, positive relation between WST performance and RMET accuracy was expected. For all analyses, data fulfilled the requirements for parametric statistical testing. Statistical analyses were performed using IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp, 2013). All graphs were created using Microsoft Excel (Microsoft, 2010) and Adobe Illustrator CS5 (Adobe Systems, 2010).

Analyses of Potentially Modulating Factors: Cognitive Performance
Especially when investigating age effects on mindreading ability, performance on other cognitive domains must be considered to test the possible influence of cognitive decline on mindreading performance at old age. Therefore, scores representative of individual performance on several cognitive domains (i.e., attention, executive function, memory, and verbal fluency; Lakens, 2013) were additionally included as predictors of RMET performance in the linear regression model. To compute cognitive domain scores, neuropsychological test scores were

TABLE 1 | Demographic characteristics and descriptive results of the Reading the Mind in the Eyes Test (RMET) and the vocabulary test (Wortschatztest, WST) for the whole sample (N = 1,603).

		Gender		Independent sample t-test t(df;n), p-value
		Men	**Women**	
n		905	698	
Age	M (SD)	60.26 (14.99)	60.35 (14)	$t(1;1544) = -0.118$, $p > 0.05$
WST (% correct)	M (SD)	78.48 (10.20)	74.92 (10.48)	$t(1;1601) = 6.805$, $p < 0.001$
RMET (% correct)	M (SD)	62.76 (10.54)	63.79 (10.51)	$t(1;1601) = -1.956$, $p = 0.051$

Note: df, degrees of freedom; M, mean; n, number; and SD standard deviation. Number of subjects and gender distribution per age group are 25–29 years 34 men/17 women, 30–34 39/17, 35–39 33/23, 40–44 57/52, 45–49 61/52, 50–54 35/27, 55–59 33/32, 60–64 105/122, 65–69 191/138, 70–74 211/149, and 75–80 90/54.

standardized to the mean of age- and gender-specific groups from the actual sample. Measures pertaining to a specific cognitive domain were further averaged to compound scores reflecting domain-specific cognitive performance (for a full description of included tests and procedures see Kynast et al., 2018). Analysis included all individuals of the current sample with full recordings of cognitive measures ($N = 1,157$; 698 men). For all analyses, data fulfilled the requirements for parametric statistical testing.

Analyses of Potentially Modulating Factors: Neuropsychiatric Diseases

The inclusion criteria for the analysis sample were broadly defined, as we aimed at investigating effects in a generic population. However, neurological diseases and psychiatric disorders potentially influence RMET accuracy besides age, gender and cognitive abilities. Accordingly, we re-tested effects in a selected subsample of the study cohort without such diseases/disorders. Individuals fulfilling one or more of the following criteria were excluded from these analyses: (a) history of neurological or psychiatric disorder, alcohol and substance abuse, dementia, (b) stroke, tumor or brain injury according to structural magnetic resonance imaging, (c) intake of psychoactive medication, (d) depression score > 20 in the Center of Epidemiologic Studies Rating Scale Depression (Radloff, 1977), and (e) extensive white matter hyperintensities exceeding Fazekas stage 1 as indices for small vessel disease (Fazekas et al., 1987; Kynast et al., 2018). Furthermore, subjects with severe cognitive dysfunction, i.e., dementia (operationalized as domain specific test performance below -2 standard deviation of the age group mean, DSM-5; American Psychiatric Association, 2013) were excluded. All analyses have been re-calculated within this selected subsample ($N = 978$; 576 men; age $M = 58.13$ years), including

(1) linear regression analysis predicting RMET performance from age, gender, age*gender and WST performance, (2) linear regression analysis predicting RMET performance from age, gender, age*gender, attention, executive function, memory, verbal fluency, and WST performance.
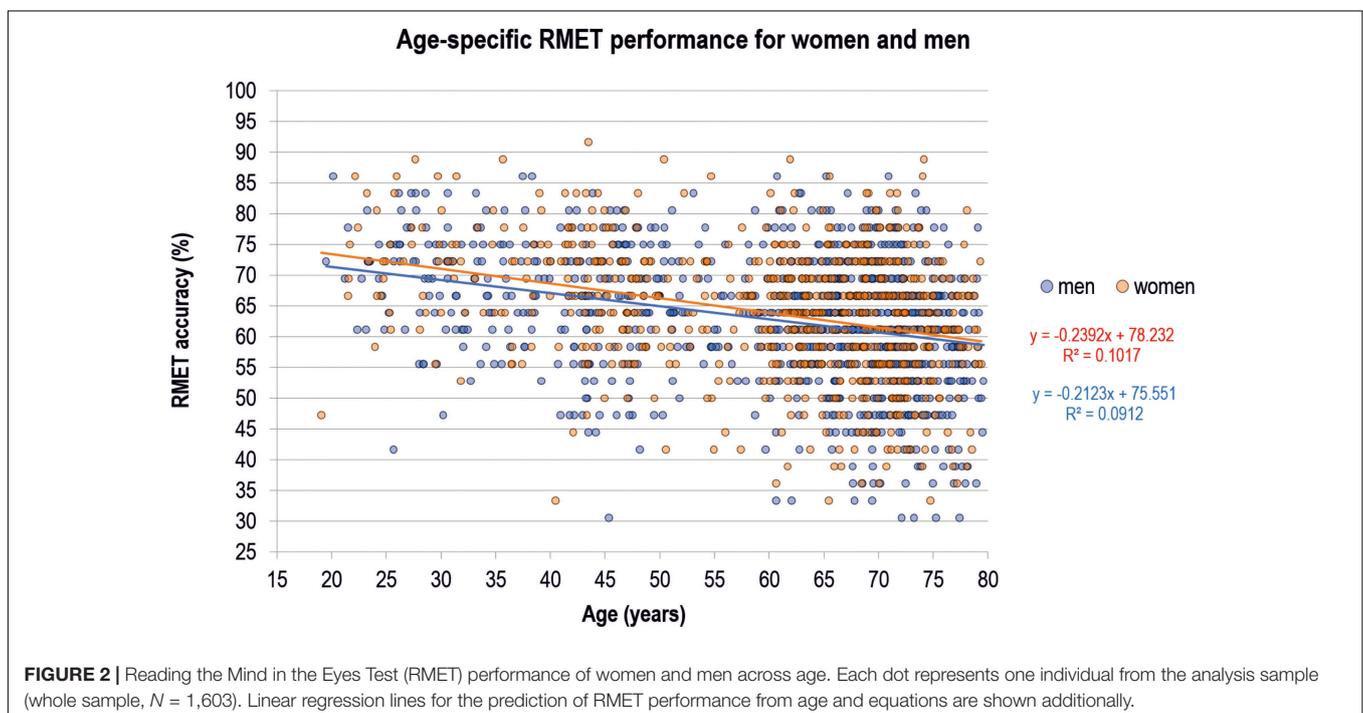
RESULTS

Test Performance

Figure 2 shows individual RMET performance scores for women and men in relation to age for the whole sample. Performance declined with aging in both women and men on a descriptive level. Mean age, WST performance reflecting verbal intelligence, and RMET performance in men compared to performance in women are listed in **Table 1**. Men scored significantly higher in the WST compared to women ($p < 0.001$), whereas the effect of gender on RMET performance just missed the significance threshold ($p = 0.051$).

Linear Regression Analyses

The linear regression model including age, gender, age*gender, and WST performance as predictors of RMET performance was significant [$F(4;1598) = 78.142$, $p < 0.001$] [results are reported as F (degrees of freedom; number of subjects)]. Older age ($b = -0.208$, $p < 0.001$) and lower verbal intelligence ($b = 0.261$, $p < 0.001$) were significantly related to lower RMET accuracy. No significant results were obtained for gender ($b = 2.677$, $p = 0.201$) and the interaction term gender*age ($b = -0.011$, $p = 0.734$). The statistical model explained 16.4% of variance in RMET accuracy ($R^2 = 0.164$).

FIGURE 2 | Reading the Mind in the Eyes Test (RMET) performance of women and men across age. Each dot represents one individual from the analysis sample (whole sample, $N = 1,603$). Linear regression lines for the prediction of RMET performance from age and equations are shown additionally.

Analyses of Potentially Modulating Factors: Cognitive Performance

Results for this analysis are illustrated in **Table 2**. The linear regression model including age, gender, age*gender, attention, executive functions, memory, verbal fluency, and WST performance as predictors of RMET performance was significant [$F(8;1148) = 30.088$, $p < 0.001$, and $R^2 = 0.145$]. Besides age ($p < 0.001$) and WST performance ($p < 0.001$), memory ($p = 0.037$) was a significant predictor of RMET performance (**Table 2**).

Analyses of Potentially Modulating Factors: Neuropsychiatric Diseases

Linear regression analyses performed on the selected subsample without neurological diseases or psychiatric disorders ($N = 978$) revealed generally comparable results to analyses performed on the entire sample. Respective results are presented in **Figure 3** and **Tables 3**, **4**. Accordingly, results of the first analysis were confirmed and were not related to possibly confounding factors such as neurological diseases or psychiatric disorders. Whereas the association between RMET performance and age as well as verbal intelligence was replicated in the linear regression analysis, the association with memory disappeared.

DISCUSSION

The RMET is one of the most frequently used tools for the assessment of ToM in adults. The primary aim of our study was testing the effects of age on RMET accuracy in a lifespan perspective. Therefore, we investigated mindreading from the eyes in a population-based lifespan sample including more than 1,600 randomly selected individuals. Crucially, we considered the potential impact of distinct cognitive functions on the association between the ability to identify mental states from gaze, and age. We found a linear relation between RMET accuracy and age. Older age was associated with lower RMET accuracy in both men and women. This is in line with other studies that have

TABLE 2 | Linear regression analysis to predict Reading the Mind in the Eyes Test (RMET) performance from age, gender, and domain-specific cognitive scores (whole sample, $N = 1,157$).

	B	**SE**	**b**	**p**
Age	−0.214	0.024	−0.297	< 0.001
Gender	2.957	2.274	0.137	0.232
Age*Gender	−0.005	0.04	−0.015	0.897
WST	0.217	0.03	0.214	< 0.001
Attention	0.710	0.377	0.053	0.06
Executive function	−0.093	0.412	−0.006	0.822
Memory	0.914	0.438	0.062	0.037
Verbal fluency	0.306	0.358	0.026	0.393

Note: B, regression coefficient (non-standardized); SE, standard error; b, regression coefficient (standardized); and WST, Wortschatztest (vocabulary test). The statistical model explains 14.5% of the variance in RMET performance corresponding to R^2 of 0.145 ($p < 0.001$).

investigated this issue. In sum, our results indicate an age-related decline in RMET performance.

It has been discussed previously if changes of ToM at older age may be specifically related to stimuli and modalities (Henry et al., 2013). Thus, the age-related performance drop in mindreading from the eyes may be related to test-specific characteristics. The RMET is an advanced test of ToM that focuses on the eyes as social cues. Thus, a decrease in test accuracy at older age may potentially be driven by a specific impairment in extracting mental state information when social cues are limited in general, and/or this limitation pertains specifically to the eyes. Older adults may prefer other social cues, such as the mouth, body posture, or voice (Phillips et al., 2014), or rely on the integration of information from multiple social sources (Hunter et al., 2010).

Additional aspects that must be considered in this context include stimulus characteristics. Individual features of the interaction partner, such as age, gender or ethnicity, have been shown to influence social information perception and processing, and behavior (Seidel et al., 2010; Slessor et al., 2010). Consequently, the features of the RMET stimuli (i.e., gender, age, and emotional valence) might potentially impact test performance differently in young and older adults (Kynast and Schroeter, 2018). Thus, elderly individuals may show deficits in accurately determining a young person's mental state compared to an older person's mental state (i.e., an in-group bias). Also, emotional complexity may be an important factor, as shown in a recent study (Duval et al., 2011). There, older participants performed as accurate as the young group when confronted with more basic emotions in the RMET but had specific difficulties with complex emotions. The "positivity effect," i.e., that older adults show a relative preference for positive over negative material in cognitive tasks (Kennedy et al., 2004; Phillips et al., 2014), may also apply to the RMET and explain lower accuracy rates compared to younger adults. However, these effects need further investigation and may be subject to future studies.
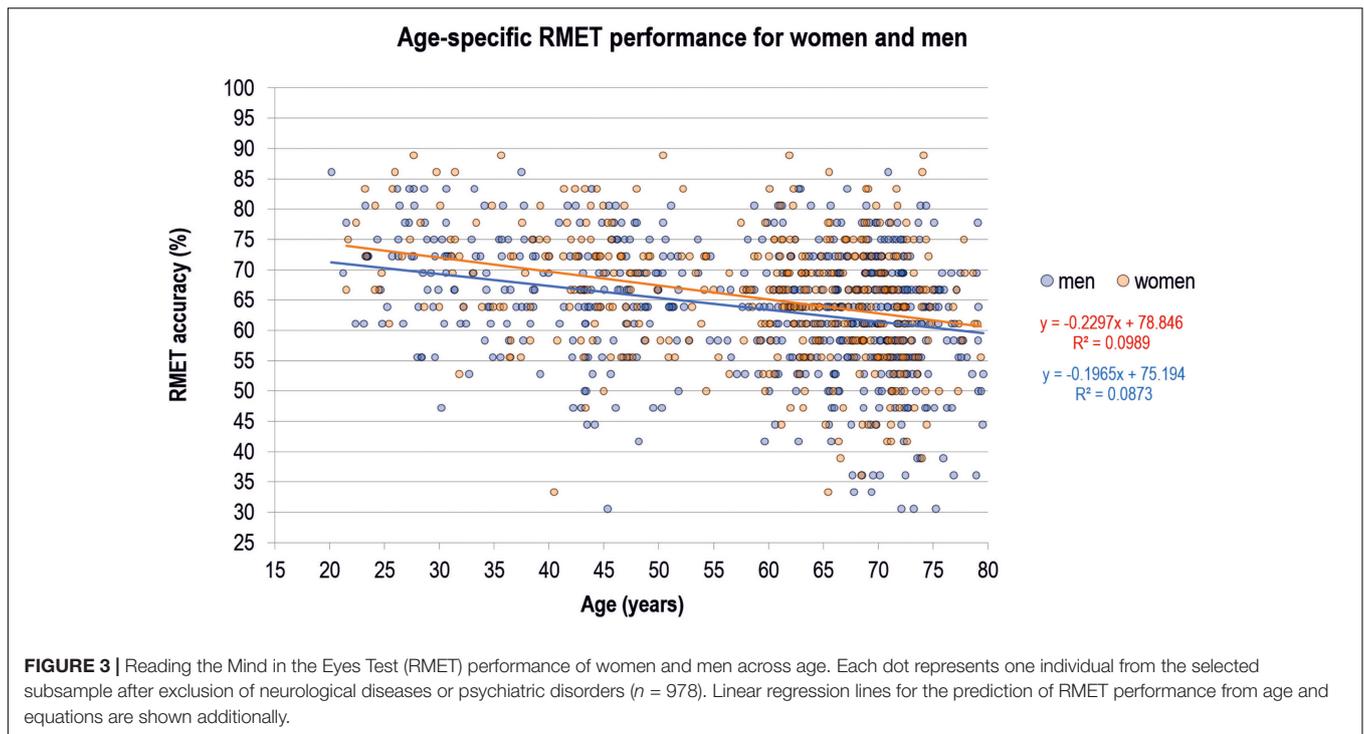
Considering cognitive abilities associated with RMET performance, a significant relation was found with both vocabulary test performance (WST) and memory. This may be best explained by the fact that these measures strongly overlap in their focus on verbal content and, thus, strongly address verbal abilities. A conceptual representation of the response alternatives given with each RMET item is required for target word selection (Baron-Cohen et al., 1997), i.e., the individual must know what the response alternatives mean to select the most appropriate term. Consequently, a comprehensive vocabulary potentially facilitates proper task completion. The results verify this assumption by indicating a relative importance of verbal abilities for mindreading from the eyes.

Although this study was carefully conducted, several limitations must be discussed. In accordance with previous studies in the field, we applied a rather conventional binary concept of sex/gender, i.e., affiliation to women or men, to characterize men's and women's ability to identify mental states from gaze. This approach, which refers rather to gender than to biological sex, enables comparability with previous studies but is a general limitation since further biological and environmental characteristics impacting individual self-concepts of sexual

FIGURE 3 | Reading the Mind in the Eyes Test (RMET) performance of women and men across age. Each dot represents one individual from the selected subsample after exclusion of neurological diseases or psychiatric disorders (n = 978). Linear regression lines for the prediction of RMET performance from age and equations are shown additionally.

identity (e.g., social roles, experiences, etc.; for review see Jäncke, 2018) were not investigated and, thus, potential effects thereof on mindreading remain still unknown. Moreover, Hyde et al. (2019) described five challenges that generally question the so called gender binary, i.e., that humans comprise only two types of beings, women and men: Neuroscience and neuroendocrinology questioning sexual brain/hormonal dimorphism (see also Jäncke et al., 2015), psychology underlining rather similarities than differences between women and men on a broad continuum (as also suggested by our data), research on transgender and non-binary individuals' identities, and research suggesting gender/sex as binary category to be culturally determined. Accordingly, Hyde (2005) postulated the gender similarities hypothesis – based on meta-analytical evidence – that males and females are similar on most psychological variables. To expand the view within social cognitive research, future studies

must incorporate a broader definition beyond the conventional concept of sex/gender to fully picture the diversity in sexual identity and its impact on cognitive processes (see Ainsworth, 2015; Dotson and Duarte, 2020). Due to the conceptual limitations of our study regarding this parameter we abstained here from discussing our results in depth.

Moreover, one may criticize that our results are potentially confounded by neurological diseases or psychiatric disorders. We can exclude such a bias, because we confirmed our results in a selected subsample, where we excluded relevant subjects. Note, that confining the cohort with this strategy also reduced the

TABLE 3 | Linear regression analysis to predict Reading the Mind in the Eyes Test (RMET) performance from age, gender, and WST performance in the selected subsample after exclusion of neurological diseases or psychiatric disorders (n = 978).

	B	*SE*	*b*	*p*
Age	−0.193	0.026	−0.279	< 0.001
Gender	3.509	2.572	0.167	0.173
Age*Gender	−0.018	0.043	−0.053	0.671
WST	0.224	0.032	0.215	< 0.001

Note. B, regression coefficient (non-standardized); SE, standard error; b, regression coefficient (standardized); and WST, Wortschatztest (vocabulary test). The statistical model explains 13.8% of the variance in RMET performance corresponding to R^2 of 0.138 (p < 0.001).

TABLE 4 | Linear regression analysis to predict Reading the Mind in the Eyes Test (RMET) performance from age, gender, and domain-specific cognitive scores in the selected subsample after exclusion of neurological diseases or psychiatric disorders (n = 978).

	B	*SE*	*b*	*p*
Age	−0.191	0.026	−0.28	< 0.001
Gender	3.693	2.646	0.714	0.163
Age*Gender	−0.018	0.044	−0.415	0.678
WST	0.204	0.035	0.192	< 0.001
Attention	0.414	0.436	0.031	0.342
Executive function	0	0.464	0	1
Memory	0.456	0.476	0.032	0.338
Verbal fluency	0.412	0.394	0.035	0.296

Note: B, regression coefficient (non-standardized); SE, standard error; b, regression coefficient (standardized); and WST, Wortschatztest (vocabulary test). The statistical model explains 13.8% of the variance in RMET performance corresponding to R^2 of 0.138 (p < 0.001).

relatively high variability in the uncontrolled cohort as shown in **Figures 2**, **3**. Furthermore, the number of participants was not balanced across the different age groups, which can be regarded statistically less-than-ideal. This difference is related to the conceptual development of the LIFE study with a main focus on the elderly age ranges. However, numbers in younger age groups are still sufficient and large if compared to earlier studies on this issue, justifying regarding results as reliable. Requirements for the application of parametric statistical methods were tested carefully. Results of our cross-sectional study shall be proven in future longitudinal studies also involving other languages and ethnicities beside our German version applied in a Caucasian cohort, and including a comprehensive conceptualizing of sex/gender.

In summary, our study investigated the effects of age, and domain-specific cognitive abilities on mental state recognition from gaze in the RMET in a population-based sample including 1,603 adults aged 19–79 years, i.e., across the whole lifespan. Results reveal declining performance in the RMET with aging for both, women and men. Taken together, ToM, and particularly mindreading from the eyes might be interpreted as a universal human feature present in both men and women over the whole lifespan but declining with aging.

DATA AVAILABILITY STATEMENT

Data are available from the authors upon reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the University of Leipzig.

The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JK planned and conducted statistical analyses, wrote the first draft of the manuscript including creation of the figures, and modified all subsequent drafts. EQ planned the study, assessed and processed data, and substantially contributed to all drafts of the manuscript. MP was responsible for data quality management, and contributed substantially to all drafts of the manuscript. TL, SR-H, and AV designed the study, interpreted results, and reviewed the final draft of the manuscript. SB-C, AH, JS, and AW contributed substantially to the interpretation of the results. MS designed the study, supervised data acquisition and analyses, substantially contributed to the interpretation of the results and made substantial modifications to all drafts of the manuscript. Further, he took the lead in revising the manuscript. All authors approved the final version of the manuscript.

FUNDING

REFERENCES

Adobe Systems (2010). *Adobe Illustrator CS5 (Version 15.1.1) [Computer Software]*. San Jose, CA: Adobe Systems.

Ahmed, F. S., and Miller, L. S. (2011). Executive function mechanisms of theory of mind. *J. Autism Develop. Disord.* 41, 667–678. doi: 10.1007/s10803-010-1087-7

Ainsworth, C. (2015). Sex redefined. *Nature* 518, 288–291. doi: 10.1038/518288a

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5 Edn. Arlington, VA: American Psychiatric Publishing. doi: 10.1176/appi.books.9780890425596

Baglio, F., Castelli, I., Alberoni, M., Blasi, V., Griffanti, L., Falini, A., et al. (2012). Theory of mind in amnestic mild cognitive impairment: an fMRI study. *J. Alzheimers Dis.* 29, 25–37. doi: 10.3233/jad-2011-111256

Bailey, P. E., and Henry, J. D. (2008). Growing less empathic with age: disinhibition of the self-perspective. *J. Gerontol. Series B Psychol. Sci. Social Sci.* 63, 219–226. doi: 10.1093/geronb/63.4.P219

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

Baron-Cohen, S., Bowen, D. C., Holt, R. J., Allison, C., Auyeung, B., Lombardo, M. V., et al. (2015). The "Reading the Mind in the Eyes" Test: complete absence of typical sex difference in ∼400 men and women with autism. *PLoS One* 10:e0136521. doi: 10.1371/journal.pone.0136521

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or Asperger syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi: 10.1111/j.1469-7610.1997.tb01599.x

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a theory of mind. *Cognition* 21, 37–46. doi: 10.1016/0010-0277(85)90 022-8

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and 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. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1017/s002196300100 6643

Bölte, S. (2005). *Reading the Mind in the Eyes Test für Erwachsene (dt. Fassung) Von S. Baron-Cohen. J.W.* Germany: Goethe-Universität.

Brune, M., and Brune-Cohrs, U. (2006). Theory of mind - Evolution, ontogeny, brain mechanisms and psychopathology. *Neurosci. Biobehav. Rev.* 30, 437–455. doi: 10.1016/j.neubiorev.2005.08.001

Dal Monte, O., Schintu, S., Pardini, M., Berti, A., Wassermann, E. M., Grafman, J., et al. (2014). The left inferior frontal gyrus is crucial for reading the mind in the eyes: brain lesion evidence. *Cortex* 58, 9–17. doi: 10.1016/j.cortex.2014.05.002

Dotson, V. M., and Duarte, A. (2020). The importance of diversity in cognitive neuroscience. *Ann. N. Y. Acad. Sci.* 1464, 181–191. doi: 10.1111/nyas.14268

Duval, C., Piolino, P., Bejanin, A., Eustache, F., and Desgranges, B. (2011). Age effects on different components of theory of mind. *Conscious. Cogn.* 20, 627–642. doi: 10.1016/j.concog.2010.10.025

El Haj, M., Raffard, S., and Gely-Nargeot, M. C. (2016). Destination memory and cognitive theory of mind in normal ageing. *Memory* 24, 526–534. doi: 10.1080/09658211.2015.1021257

Fazekas, F., Chawluk, J. B., Alavi, A., Hurtig, H. I., and Zimmerman, R. A. (1987). MR signal abnormalities at 1.5-T in Alzheimer dementia and normal aging. *Am. J. Roentgenol.* 149, 351–356. doi: 10.2214/ajr.149.2.351

Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S., et al. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. *Brain* 125, 752–764. doi: 10.1093/brain/awf079

Henry, J. D., Phillips, L. H., Ruffman, T., and Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychol. Aging* 28, 826–839. doi: 10.1037/a0030677

Hunter, E. M., Phillips, L. H., and MacPherson, S. E. (2010). Effects of age on cross-modal emotion perception. *Psychol. Aging* 25, 779–787. doi: 10.1037/a0020528

Hyde, J. S. (2005). The gender similarities hypothesis. *Am. Psychol.* 60, 581–592. doi: 10.1037/0003-066X.60.6.581

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., and van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *Am. Psychol.* 74, 171–193. doi: 10.1037/amp0000307

IBM Corp (2013). *IBM SPSS Statistics for Windows (Version 22.0) [Computer Software]*. Armonk, NY: IBM Corp.

Jäncke, L. (2018). Sex/gender differences in cognition, neurophysiology, and neuroanatomy [version 1; peer review: 3 approved]. *F1000Research* 7:805. doi: 10.12688/f1000research.13917.1

Jäncke, L., Mérillat, S., Liem, F., and Hänggi, J. (2015). Brain size, sex, and the aging brain. *Hum. Brain Mapp.* 36, 150–169. doi: 10.1002/hbm.22619

Kennedy, Q., Mather, M., and Carstensen, L. L. (2004). The role of motivation in the age-related positivity effect in autobiographical memory. *Psychol. Sci.* 15, 208–214. doi: 10.1111/j.0956-7976.2004.01503011.x

Kirkland, R. A., Peterson, E., Baker, C. A., Miller, S., and Pulos, S. (2013). Meta-analysis reveals adult female superiority in "Reading the Mind in the Eyes". *Test. North Am. J. Psychol.* 15, 121–146.

Kynast, J., Lampe, L., Luck, T., Frisch, S., Arelin, K., Hoffmann, K. T., et al. (2018). White matter hyperintensities associated with small vessel disease impair social cognition beside attention and memory. *J. Cereb. Blood Flow Metab.* 38, 996–1009. doi: 10.1177/0271678x17719380

Kynast, J., and Schroeter, M. L. (2018). Sex, age, and emotional valence: revealing possible biases in the 'Reading the Mind in the Eyes' task. *Front. Psychol.* 9:570. doi: 10.3389/fpsyg.2018.00570

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* 4:863. doi: 10.3389/fpsyg.2013.00863

Loeffler, M., Engel, C., Ahnert, P., Alfermann, D., Arelin, K., Baber, R., et al. (2015). The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany. *BMC Public Health* 15:691. doi: 10.1186/s12889-015-1983-z

Microsoft (2010). *Microsoft Excel 2010 [Computer Software]*. Redmond, WA: Microsoft.

Morton, J., and Johnson, M. H. (1991). Conspec and conlern -A 2-process theory of infant face recognition. *Psychol. Rev.* 98, 164–181. doi: 10.1037//0033-295x.98.2.164

Pardini, M., Gialloreti, L. E., Mascolo, M., Benassi, F., Abate, L., Guida, S., et al. (2013). Isolated theory of mind deficits and risk for frontotemporal dementia: a longitudinal pilot study. *J. Neurol. Neurosurg. Psychiatry* 84, 818–821. doi: 10.1136/jnnp-2012-303684

Pardini, M., and Nichelli, P. F. (2009). Age-related decline in mentalizing skills across adult life span. *Exp. Aging Res.* 35, 98–106. doi: 10.1080/03610730802545259

Peterson, E., and Miller, S. F. (2012). The eyes test as a measure of individual differences: how much of the variance reflects verbal IQ? *Front. Psychol.* 3:220. doi: 10.3389/fpsyg.2012.00220

Phillips, L. H., Slessor, G., Bailey, P. E., and Henry, J. D. (2014). "Older adults' perception of social and emotional cues," in *The Oxford Handbook of Emotion, Social Cognition, and Everyday Problem Solving During Adulthood*, eds P. Verhaeghen, and C. Hertzog (Oxford: Oxford University Press).

Poletti, M., and Bonuccelli, U. (2013). Alteration of affective theory of mind in amnestic mild cognitive impairment. *J. Neuropsychol.* 7, 121–131. doi: 10.1111/j.1748-6653.2012.02040.x

Premack, D., and Woodruff, G. (1978). Chimpanzee problem-solving: a test for comprehension. *Science* 202, 532–535. doi: 10.1126/science.705342

Radloff, L. S. (1977). The CES-D Scale: a self-report depression scale for research in the general population. *Appl. Psychol. Meas.* 1, 385–401. doi: 10.1177/014662167700100306

Schmidt, K. H., and Metzler, P. (1992). *WST-Wortschatztest*. Weinheim: Beltz Test GmbH.

Schroeter, M. L. (2012). Considering the frontomedian cortex in revised criteria for behavioural variant frontotemporal dementia. *Brain* 135:e213. doi: 10.1093/brain/aws030

Schroeter, M. L., Laird, A. R., Chwiesko, C., Deuschl, C., Schneider, E., Bzdok, D., et al. (2014). Conceptualizing neuropsychiatric diseases with multimodal data-driven meta-analyses–The case of behavioral variant frontotemporal dementia. *Cortex* 57, 22–37. doi: 10.1016/j.cortex.2014.02.022

Schroeter, M. L., Pawelke, S., Bisenius, S., Kynast, J., Schuemberg, K., Polyakova, M., et al. (2018). A modified Reading the Mind in the Eyes Test predicts behavioral variant frontotemporal dementia better than executive function tests. *Front. Aging Neurosci.* 10:11. doi: 10.3389/fnagi.2018.00011

Schroeter, M. L., Zysset, S., Kruggel, F., and von Cramon, D. Y. (2003). Age dependency of the hemodynamic response as measured by functional near-infrared spectroscopy. *NeuroImage* 19, 555–564. doi: 10.1016/s1053-8119(03)00155-1

Seidel, E. M., Habel, U., Kirschner, M., Gur, R. C., and Derntl, B. (2010). The impact of facial emotional expressions on behavioral tendencies in women and men. *J. Exp. Psychol. Hum. Percep. Perform.* 36, 500–507. doi: 10.1037/a0018169

Slessor, G., Miles, L. K., Bull, R., and Phillips, L. H. (2010). Age-related changes in detecting happiness: discriminating between enjoyment and nonenjoyment smiles. *Psychol. Aging* 25, 246–250. doi: 10.1037/a0018248

Warrier, V., Grasby, K. L., Uzefovsky, F., Toro, R., Smith, P., Chakrabarti, B., et al. (2018). Genome-wide meta-analysis of cognitive empathy: Heritability, and correlates with sex, neuropsychiatric conditions and cognition. *Mol. Psychiatry* 23, 1402–1409. doi: 10.1038/mp.2017.122

Whiten, A. (1991). How monkeys see the world. *Curr. Anthropol.* 32, 503–505. doi: 10.1086/203992