



Age- and Sex-Specific Standard Scores for the Reading the Mind in the Eyes Test

Jana Kynast^{1,2}, Maryna Polyakova^{1,2}, Eva Maria Quinque^{1,3}, Andreas Hinz⁴, Arno Villringer^{1,2,3} and Matthias L. Schroeter^{1,2,3*}

¹ Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, ² LIFE - Leipzig Research Centre for Civilization Diseases, University of Leipzig, Leipzig, Germany, ³ Clinic for Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany, ⁴ Department of Medical Psychology and Sociology, University Hospital Leipzig, Leipzig, Germany

The reliable, valid and economic assessment of social cognition is more relevant than ever in the field of clinical psychology. Theory of Mind is one of the most important socio-cognitive abilities but standardized assessment instruments for adults are rare. The Reading the Mind in the Eyes Test (RMET) is well-established and captures the ability to identify mental states from gaze. Here, we computed standard scores for the German version of the RMET derived from a large, community-dwelling sample of healthy adults (20–79 years). The standardization sample contains 966 healthy adult individuals of the population-based Leipzig Research Center for Civilization Diseases (LIFE) study. Before standardization, weighting factors were applied to match the current sample with distribution characteristics of the German population regarding age, sex, and education. RMET scores were translated into percentage ranks for men and women of five age groups (20–29, 30–39, 40–49, 50–59, 60+ years). Age-specific percentage ranks are provided for men and women. Independent of age, men present a larger variance in test scores compared to women. Within the specific age groups, women score higher and their scoring range is less variable. With increasing age, the scoring variance increases in both men and women. This is the first study providing age- and sex-specific RMET standard scores. Data was weighted to match German population characteristics, enabling the application of standard scores across German-speaking areas. Our results contribute to the standardized assessment of socio-cognitive abilities in clinical diagnostics.

Keywords: theory of mind, mindreading, Reading the Mind in the Eyes test, social cognition, neuropsychological testing, cognitive aging, dementia

INTRODUCTION

Social cognition is one of the most complex and unresolved concepts in psychology and neurosciences. It encompasses numerous definitions (Happé et al., 2017) and includes many socio-cognitive abilities. The ability to identify mental states, intentions, desires, feelings and beliefs of self and others (Premack and Woodruff, 1978) is called Theory of Mind (ToM). It comprises the perception, processing and integration as well as the interpretation of social context information facilitating the prediction of another individual's future actions and the adaptation

OPEN ACCESS

Edited by:

Guido Gainotti,
Catholic University of the Sacred
Heart, Italy

Reviewed by:

Julie Suhr,
Ohio University, United States
Andrei C. Miu,
Babeş-Bolyai University, Romania

*Correspondence:

Matthias L. Schroeter
schroet@cbs.mpg.de

Received: 16 September 2020

Accepted: 28 December 2020

Published: 28 January 2021

Citation:

Kynast J, Polyakova M, Quinque EM, Hinz A, Villringer A and Schroeter ML (2021) Age- and Sex-Specific Standard Scores for the Reading the Mind in the Eyes Test. *Front. Aging Neurosci.* 12:607107. doi: 10.3389/fnagi.2020.607107

of own behavioral responses. Thus, ToM contributes to successful social interactions. Contrary, impairment in ToM can compromise the initiation and maintenance of positive relationships. Among various psychiatric disorders and neurological diseases, brain injury, age-related brain changes, as well as developmental and neurodegenerative disorders (e.g., Gregory et al., 2002; Brune and Brune-Cohrs, 2006; Schroeter et al., 2010, 2014; Schroeter, 2012; Kynast et al., 2018; Vallat-Azouvi et al., 2018) have been characterized by ToM deficits. Further, the clinical relevance of ToM increased recently, when the Diagnostic Statistical Manual of Mental Disorders' 5th edition (DSM-5; American Psychiatric Association, 2013) explicitly defined social cognition as one key domain for neuropsychological assessment, especially in the context of acquired cognitive dysfunction, i.e., major and mild neurocognitive disorder (NCD; corresponding terms are dementia and its prestage, i.e., mild cognitive impairment, MCI).

Standardized psychological tests help to identify specific deficits (and resources) as they provide normative scores for the classification of individual abilities relative to a reference group (ideally) age-, sex- and education-matched. Performance deviations might indicate the need for subsequent in-depth testing to further evaluate and confirm the detected deficits as well as to classify potential impairment regarding activities of daily living. However, the systematic assessment of socio-cognitive abilities, particularly ToM, is problematic, as only a few instruments are available for adults and even fewer are standardized.

The *Reading the Mind in the Eyes Test* (RMET; Baron-Cohen et al., 1997, 2001) may be a potential candidate for standardized assessment of specific aspects of social cognition. It captures the ability to identify mental states based on the eyes of another person. The revised version (Baron-Cohen et al., 2001) includes 36 photographs of the eye region of 18 men and 18 women. Each item comes along with four response options. The term most appropriately describing the mental state of the pictured person shall be selected. The RMET requires the identification of complex mental states where social cues are limited (Baron-Cohen et al., 1997). The task addresses especially the identification of the relevant mental state of the stimulus (first stage of ToM attribution) rather than inferring the content of that mental state (second stage; see Baron-Cohen et al., 2001). It is supposed to capture predominantly "hot" or "affective" ToM, since it requires an understanding of another person's affective states or feelings (c.f. "cold" or "cognitive" ToM: a concept of cognitive states, beliefs, thoughts, or intentions; Henry et al., 2013; El Haj et al., 2015). This might be in line with a recent critical note on the validity of the RMET (Oakley et al., 2016), suggesting that the focus of the RMET is rather emotion recognition than ToM as compared to other assessment instruments using cartoons or short movies of social situations. Overall, studies show mixed results regarding the validity of the test (see Olderbak et al., 2015 for review of the psychometric properties; critical aspects regarding test construction can be found in Kynast and Schroeter, 2018). Besides these critical points, the RMET may still be a promising candidate for further clinical application since it

has been shown to identify deficits in mental state recognition in psychopathology and neurodegenerative diseases such as behavioral variant frontotemporal dementia (Pardini et al., 2013; Schroeter et al., 2018), Alzheimer's disease and mild NCD (see Baglio et al., 2012), autism (Baron-Cohen et al., 1997, 2015), bipolar disorder (Bora et al., 2005) and small vessel disease (Kynast et al., 2018). The test is not characterized by a ceiling effect since the response rate of healthy individuals is typically below 100% (Pardini and Nichelli, 2009). Furthermore, it is one of the most frequently used tasks to investigate ToM in adulthood, published in more than 250 (Kirkland et al., 2013) or even 542 studies according to PubMed until 6th of April 2020 (term = reading + the + mind + in + the + eyes). Test performance is modulated by individual characteristics such as sex, and age, and the characteristics of the stimulus material itself (e.g., Baron-Cohen et al., 1997; Bailey et al., 2008; Pardini and Nichelli, 2009; Castelli et al., 2010; Cabinio et al., 2015; El Haj et al., 2015; Fischer et al., 2016; Kynast and Schroeter, 2018; Kynast et al., 2018, 2020). To our best knowledge, no standard scores are yet available in any language.

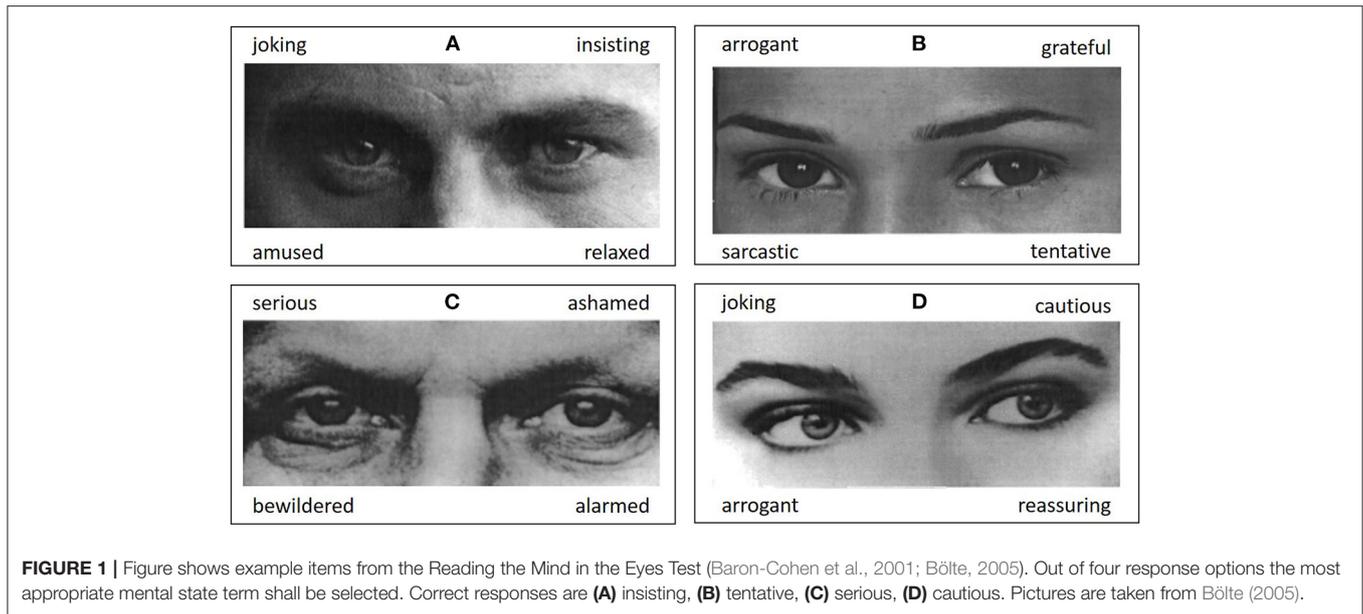
This study is the first to present age-standardized scores for the RMET derived from a large, population-based sample of healthy adults (Loeffler et al., 2015). We used the German version of the RMET (Bölte, 2005). Weighting factors were applied to match the current sample to the characteristic distribution of age, sex and education represented in the German population. RMET scores were transformed into percentile ranks that characterize the performance of an individual in relation to its reference group. The study aimed overall at yielding norms for neuropsychological characterization of adults and potentially identifying deviations due to disease.

METHODS

Study Cohort

This study was part of the adult study of the Leipzig Research Center for Civilization Diseases (LIFE). The LIFE adult study is a cross-sectional study investigating "prevalences, early onset markers, genetic predispositions, and the role of lifestyle factors of major civilization diseases" (Loeffler et al., 2015). The study was approved by the ethics committees of the University of Leipzig and was conducted in accordance with the latest version of the Declaration of Helsinki. Each subject provided written informed consent. All participants were randomly selected residents of the city of Leipzig (Saxony/Germany, population ~600,000, Amt für Statistik und Wahlen Leipzig, 2020) aged 18–79 years. Participation was voluntary. The total sample comprised 10,000 individuals of whom 2,600 completed structural and functional brain magnetic resonance imaging (MRI). Adults aged 60–79 years completed an in-depth-assessment (neuropsychological testing, medical examinations, interviews on individual lifestyle conditions). Persons younger than 60 years completed a less extensive test battery. The complete protocol is described elsewhere (Loeffler et al., 2015).

For the current study, participants who fulfilled at least one of the following criteria were excluded from data analysis: (1) history of neurological or psychiatric disorder (i.e., dementia,



alcohol or substance abuse, schizophrenia, affective and anxiety disorders, eating disorders, autism), (2) intake of medication active on the central nervous system (opioids, hypnotics and sedatives, anti-parkinsonian drugs, anxiolytics, antipsychotics, anti-epileptic drugs), (3) depression score >20 on the Center of Epidemiologic Studies Depression Scale (Radloff, 1977), (4) history of stroke, brain injury or tumor, (5) significant white matter hyperintensities (Fazekas stages 2 and 3) on T2-weighted fluid attenuated inversion recovery (FLAIR) MRI scans (Fazekas et al., 1987). The latter has been selected as a specific exclusion criterion since age-related changes in the brain's white matter structure are associated with cognitive impairment. Especially structural alterations as severe as Fazekas stages 2 and 3 have been associated with significant cognitive deficits in attention, memory, and most crucially, social cognition compared to individuals without white matter hyperintensities (Kynast et al., 2018).

The final standardization sample included 966 adults (48.6% male) aged 20–79 years (mean, $M = 50.7$; standard deviation, $SD = 16.2$). Demographic characteristics, self-report on medical history and medication intake, as well as brain MRI was available for the whole sample. Note that all individuals had normal or corrected to normal vision.

Test and Procedure

The ability to identify mental states from gaze was assessed with the German version of the revised RMET (Baron-Cohen et al., 2001; Bölte, 2005). The test contains 36 black-and-white photographs of the eye region of either a man or a woman. Each item is presented with four response options of which the word best describing the pictured mental state shall be selected (see Figure 1 for examples). Computerized assessment included standard instructions and a test item (Bölte, 2005). If necessary, trained study assistants provided help with the handling of the hardware for response selection. Assessment was self-paced and needed approximately 10–15 min

for completion. Prior to analysis, data was carefully checked for plausibility, i.e., possible biases in response profiles as well as accuracy rates below chance level (<25%), indicating potential nonconformity with the test instructions. No individual of the final analysis sample scored below chance level nor did study documentation indicate irregularities regarding task completion.

Statistical Analysis

Dataset Preparation

Firstly, the analysis sample was stratified into sex (male/female), age [20–29 years, 30–39 years, 40–49 years, 50–59 years, and 60 years and older (60+)] and education (≤ 10 years of school; >10 years of school) groups providing a basic data structure for the calculation of RMET standard scores. Since these standard scores should be applicable for German-speaking subjects, it was necessary that the distribution characteristics of the study sample match the German population regarding the above-mentioned characteristics. As this was not the case *per se*, weighting factors were computed based on German population data published in the “Mikrozensus 2014” study (Statistisches Bundesamt Wiesbaden, 2018). The weighting factors were derived by aligning the number of individuals of the Mikrozensus 2014 age-sex-education subsamples with the number of individuals of the corresponding subsamples of the current study. Weighting factors were then multiplied group-wise with the number of individuals of the study dataset resulting in an overlap with the German population characteristics. See Table 1 for a detailed overview about weighting factor extraction and application to the study dataset.

RMET Standard Score Computation

This computation is performed on the aligned dataset (i.e., where weighting factors were applied to match the German population according to age, sex, education). Age-specific frequency distributions of correct responses in the RMET for

TABLE 1 | Number of participants per group in the German reference population (Mikrozensus 2014; $N = 63,292$), the LIFE adult study ($N = 966$), and the weighted analyses sample (LIFE*, $N = 966$).

Sex	Education	N Mikrozensus (%)	N LIFE (%)	Weighting factor	N LIFE*
A. Age group: 20–29 years					
Male	≤10 y	2437 (3.9)	12 (1.2)	3.1	37
	>10 y	2156 (3.4)	30 (3.1)	1.1	33
Female	≤10 y	2043 (3.2)	3 (0.3)	10.4	31
	>10 y	2388 (3.8)	15 (1.6)	2.4	36
Total	≤10 y	4480 (7.1)	15 (1.6)	4.6	68
	>10 y	4544 (7.2)	45 (4.7)	1.5	69
B. Age group: 30–39 years					
Male	≤10 y	2711 (4.3)	20 (2.1)	2.1	41
	>10 y	2038 (3.2)	38 (4)	0.8	31
Female	≤10 y	2518 (4)	9 (0.9)	4.3	38
	>10 y	2186 (3.5)	24 (2.5)	1.4	33
Total	≤10 y	5229 (8.3)	29 (3)	2.8	80
	>10 y	4224 (6.7)	62 (6.4)	1	64
C. Age group: 40–49 years					
Male	≤10 y	3809 (6)	58 (6)	1	58
	>10 y	2076 (3.3)	27 (2.8)	1.2	32
Female	≤10 y	3800 (6)	44 (4.6)	1.3	58
	>10 y	1908 (3)	20 (2.1)	1.5	29
Total	≤10 y	7609 (12)	102 (10.6)	1.1	116
	>10 y	3984 (6.3)	47 (4.9)	1.3	61
D. Age group: 50–59 years					
Male	≤10 y	4075 (6.4)	28 (2.9)	2.2	62
	>10 y	1800 (2.8)	18 (1.9)	1.5	27
Female	≤10 y	4376 (6.9)	27 (2.8)	2.5	67
	>10 y	1556 (2.5)	9 (0.9)	2.6	24
Total	≤10 y	8451 (13.4)	55 (5.7)	2.3	129
	>10 y	3356 (5.3)	27 (2.8)	1.9	51
E. Age group: ≥60 years					
Male	≤10 y	7281 (11.5)	201 (20.8)	0.6	111
	>10 y	2349 (3.7)	138 (14.3)	0.3	36
Female	≤10 y	10351 (16.4)	176 (18.2)	0.9	158
	>10 y	1434 (2.3)	69 (7.1)	0.3	22
Total	≤10 y	17632 (27.9)	377 (39)	0.7	269
	>10 y	3783 (6)	207 (21.4)	0.3	58

Panels A–E show group-specific absolute number of participants per group in the German reference population (Mikrozensus 2014) and the LIFE adult study. Relative numbers are presented in parenthesis. Weighting factors are calculated as: relative number (Mikrozensus)/relative number (LIFE). In a last step, weighting factors are multiplied with the original number of participants from the LIFE adult study resulting in the final distribution of the current study's analysis sample (N LIFE*).

men and women have been used. For each frequency distribution, percentile ranks were computed as $PR_v = \frac{freq_{cum}(x_v)}{N} * 100$ (Moosbrugger and Kelava, 2012, p. 176 et seqq.). The term $freq_{cum}(x_v)$ defines the accumulated frequencies up to (and including) the score of interest (i.e., RMET accuracy score); N defines the number of participants within the (sub-) sample. The percentage rank informs about an individual's accuracy in the RMET within the reference group. It is defined as the relative number (percentage) of scores within the frequency distribution equal to or lower than it. Percentile ranks were computed separately for men and women of five age groups (20–29, 30–39, 40–49, 50–59, 60+).

RESULTS

The sample comprises 966 adults (469 male) aged 20–79 ($M = 50.7$, $SD = 16.2$; see **Table 2**). Descriptive results suggest a trend toward an inverse relation between RMET accuracy rates and age. Also, females scored slightly higher than males on a descriptive level. The sex-specific RMET raw score distribution across the sample is additionally presented in **Figure 2**. Although women scored on the top level (max accuracy score = 32), both men and women presented identical accuracy spans (20 scores).

Persons with higher education presented better RMET performance compared to persons with <10 years of school

TABLE 2 | Sociodemographic characteristics of the standardization sample ($N = 966$) and corresponding RMET scores (M , SD).

		<i>N</i>	%	RMET total correct <i>M</i> (<i>SD</i>)
Age (years)	$M = 50.7$	966	100	
	$SD = 16.2$ range: 20–79			
Age groups	20–29	138	14.2	26.0 (3.2)
	30–39	144	14.9	24.7 (3.2)
	40–49	177	18.3	24.1 (3.3)
	50–59	180	18.7	23.4 (3.2)
	60+	327	33.8	22.4 (3.8)
Sex	Male	469	48.6	23.4 (3.6)
	Female	497	51.4	24.1 (3.7)
Education	≤10 years	662	68.6	23.3 (3.7)
	>10 years	304	31.4	24.8 (3.4)

N, sample size; *M*, mean; *RMET*, Reading the Mind in the Eyes Test; *SD*, standard deviation.

(**Table 2**). An one-way ANOVA predicting RMET accuracy from age-group (5), sex (2) and education (2) was computed [$F_{(17;945)} = 9.992$, $p < 0.001$; $R^2=13.6$]. Main effects of the factors age [$F_{(4;945)} = 24.452$, $p < 0.001$], sex [$F_{(1;945)} = 8.727$, $p < 0.001$], and education [$F_{(1;945)} = 6.571$, $p < 0.001$] were found.

Normative scores are shown in **Table 3**. Here, ranked raw test scores were monotonously transformed into percentile ranks within a specific age and sex group (see again **Table 1** for information on weighting factor extraction and application to the study dataset). The maximum scores within an age group are comparable between men and women. Regarding age, older age groups present a wider scoring range compared to younger age groups. For older men/women, lower accuracy rates may therefore be defined as age-appropriate, while an identical score may be classified as “below average” in a younger individual. For instance, a number of 20 correctly identified items corresponds to a percentile rank of 12 in a man younger than 30 (below average), while performance is on average for a man in his 60’s with an identical score. Note, that for certain raw scores no corresponding percentage rank could be calculated.

DISCUSSION

In this study, we computed age- and sex-standardized scores for the German version of the RMET based on a population-based sample including 966 healthy adults. Scores were weighted to match the demographic characteristics of the German population. This is the first study publishing standard scores for the frequently used RMET. The RMET assesses a specific aspect of ToM, i.e., the ability to identify complex mental states from gaze. Based on the guidelines for the evaluation of cognitive performance in the framework of NCD introduced by the DSM-5 (American Psychiatric Association, 2013), we provide age- and sex-specific standard scores improving the reliable distinction between typical and atypical test performance for a potential application in the clinical diagnostics of social cognitive abilities.

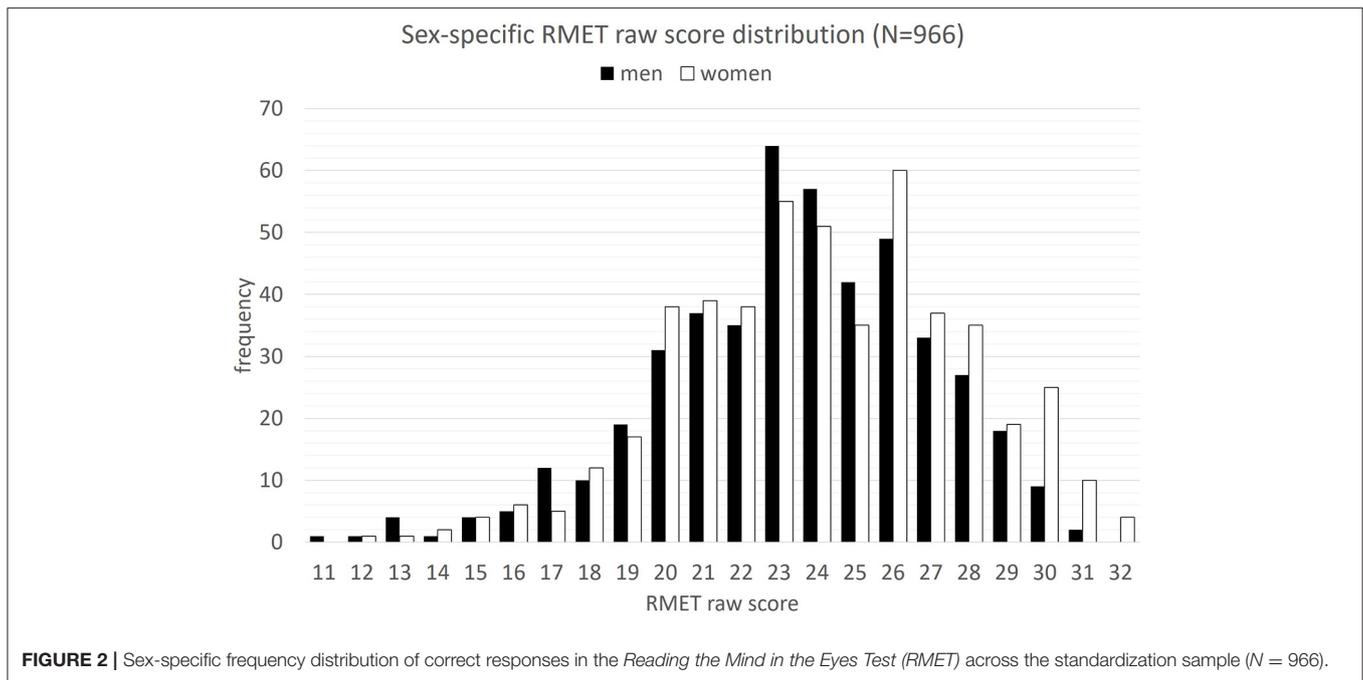
Notably, this study goes beyond clinical aspects of socio-cognitive impairment, as it addresses the ability to identify mental states from gaze in healthy adults across the adult age range. With this, it is linked to the concept of lifespan development (Baltes, 1987). Age may affect social cognitive abilities differently in men and women, but recent findings regarding the RMET are contradictory (e.g., Kirkland et al., 2013). Our results suggest slight performance differences between men and women, justifying the computation of specific standard scores. In both men and women, performance is reduced with older age.

Besides the remarkably large number of individuals used for standardization and the wide age-span covered, data was weighted according to age, sex, and education status to match the population characteristics of Germany. With this, standard scores are not limited to the current sample, but may be generally applied across German-speaking areas. Notably, rank transformation does not require normal distribution, but the shape of the distribution function informs about the group-specific frequency of RMET scores (cf. Moosbrugger and Kelava, 2012). This may be helpful for the interpretation of individual performance in relation to the standardization sample. Importantly, sections within the frequency distribution where scores are denser may lead to an overestimation of performance differences, while performance differences in less dense sections may be underestimated.

By addressing these points, our study contributes essentially to a systematic assessment of socio-cognitive abilities, which is an important aspect in clinical diagnostics of psychiatric, neurological, and neurodegenerative diseases associated with ToM impairment. Moreover, defining the range of extraordinary test performance enables to identify persons with distinct social cognitive abilities, which may serve as an important personal resource and cognitive reserve deferring cognitive decline.

Limitations and Implications for Application in Diagnostic Settings

Although this study was carefully conducted, some limitations need to be considered. Firstly, the stratification of the current sample according to sex, age, and education as a basic structure for standard score computation may seem arbitrary. However, it is common sense in the field of neuropsychology that basic demographic factors potentially impact cognitive performance. The ANOVA results additionally justify this approach. Additionally, it enables comparability with other studies and facilitates application in clinical settings. The binary coding of sex/gender information may be regarded conventional as it neglects persons with differing sexual identity and personal self-concept potentially influencing socio-cognitive functioning (e.g., Kung, 2020). Thus, standard scores provided here may not be applicable in these cases. In future, it may be helpful to consider concepts beyond the classical sex/gender dimorphism picturing sexual diversity in populations. Of note, information on sex/gender was obtained directly from the resident’s registration office of the city of Leipzig, and, additionally, via self-report. All participants showed agreement



on both factors (see also Kynast et al., 2020). Furthermore, other factors such as environment, culture and experience may drive (social) cognitive performance beyond sex/gender (see Jäncke, 2018 for review; but also Dotson and Duarte, 2020 for the importance of sex/gender in neuroscience). Secondly, education was dichotomized based on years of formal education and graduation diploma for practicability reasons. Yet, tertiary education or further professional training was not considered for standardization although it possibly modulates RMET performance. This effect on mental state attribution remains subject to future research. Thirdly, numbers of subjects differ in the several groups. Here, the relatively large age group 60+ spans two decades of adults instead of being separated into one group for each decade (60–69, 70–79). This was done to provide an appropriate number of elderly individuals with higher education (cf. Table 1) for standardization, since formal scholastic education of more than 10 years was unusual among persons growing up from between the late 1930s and early 1950s in Germany, which can be regarded as a cohort effect. Although a broad definition of this age group may thus be justified and the application of weighting factors ensures matching distribution characteristics with the German population, it must be considered that the individual test performance of a person with advanced age may be characterized less precisely under these conditions. The original study distribution did not exactly match the characteristic proportions regarding age, sex and education in the German population, which is a limitation of the study. We tried to overcome this issue by the application of weighting factors. We based this on the latest available distribution data (Statistisches Bundesamt Wiesbaden, 2018). With more than 950 individuals, the sample size can be regarded appropriate for standardization. Yet, the subsamples used for

the computation of sex-specific standard scores contain <100 individuals per group and it must be noted that in those groups not all percentage ranks may correspond to distinct RMET raw scores. Hence, our data corresponds with previous studies since again, the test does not have a ceiling effect (Pardini and Nichelli, 2009), supporting its eligibility for clinical application.

However, it must be critically addressed that the standard scores provided here are not sufficient for the diagnosis of psychiatric disorders, neurological or neurodegenerative diseases but rather indicate difficulties in a specific aspect of ToM, i.e., the ability to identify mental states from gaze. Deficits obtained in this test must be verified by subsequent, in-depth diagnostic procedures. Furthermore, RMET performance may be modulated by other individual factors that have not been considered here. For instance, it has been shown that RMET performance is related to verbal intelligence, but also other cognitive abilities (Ahmed and Miller, 2011; Peterson and Miller, 2012; Baker et al., 2014; Cabinio et al., 2015; Kynast et al., 2018, 2020) that possibly enhance or decrease mindreading accuracy. Moreover, test performance may be influenced by the RMET's stimulus characteristics (Kynast and Schroeter, 2018). Thus, these factors should be additionally assessed and considered in diagnostic settings.

Overall, the RMET, with all its strengths and limitations, can thus be a useful assessment instrument indicating socio-cognitive deficits in specific psychiatric disorders and neurological diseases. In a more resource-oriented view this test may also be used to detect individual strengths in this cognitive domain that might improve the use of therapeutic interventions or be used in professional settings (e.g., identify applicants with extraordinary skills in specific sectors such as health care or elderly care).

TABLE 3 | Age- and sex-specific percentile ranks corresponding to number of correct responses in the *Reading the Mind in the Eyes Test (RMET)*.

Sex	Men					Women					
	Age group (years)	20–29	30–39	40–49	50–59	60+	20–29	30–39	40–49	50–59	60+
N		70	72	90	90	147	68	72	87	91	180
RMET raw score	Percentile rank										
≥32							98				
31							96	98			
30	98	98				99	89			97	99
29	92	94	98			98	70	96	92		97
28	84	91	94	95		97	66	82	87	94	95
27	72	88	85	91		92	47	74	80		92
26	63	77	78	82		89	40	56	74	88	85
25	46	58	72	73		82	33	47	51	69	79
24	41	50	61	63		72	29		42	63	70
23	34	37	47	53		58	19	41	32	47	58
22	23	24	32			49	15		21	41	46
21		15	25	30		38		19	16	33	36
20	12	9	20	15		29		15	10	19	26
19		6	12	12		21		5		11	17
18			10	6		15			3	5	13
17		1	8	4		11			1		9
16			3			8					7
15			1	2		6					4
14						5					2
13						4					1
≤12						1					

N, sample size.

Conclusion

In conclusion, this is the first study providing standard scores for the ability to identify mental states from gaze. They may be used for the detection of socio-cognitive deficits in clinical practice, e.g., in the context of dementia, other neurological diseases and psychiatric disorders, or the assessment of socio-cognitive resources that may be strengthened by prevention or rehabilitation strategies. This large sample of healthy adults was weighted to match the distribution characteristics regarding age, sex, and education of the German population, enabling the application of RMET standard scores to German-speaking areas.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: Data can be requested upon reasonable request. Requests to access these datasets should be directed to Matthias L. Schroeter, schroet@cbs.mpg.de.

ETHICS STATEMENT

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

patients/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. MP was responsible for data quality management and contributed substantially to all drafts of the manuscript. EQ planned the study, assessed and processed data, and substantially contributed to all drafts of the paper. AH contributed substantially to the interpretation of the results and to the manuscript. AV designed the study, interpreted results, and reviewed the final draft of the manuscript. 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. All authors contributed to the article and approved the submitted version.

FUNDING

This study has been supported by LIFE – Leipzig Research Center for Civilization Diseases, Universität Leipzig. LIFE is

funded by means of the European Union, by the European Regional Development Fund (ERDF) and by means of the Free State of Saxony within the framework of the excellence initiative. Part of this work has been supported by the German Research Foundation (DFG; SCHR 774/5-1 to MS). JK was supported by the Max Planck International Research Network on Aging, MP by the International Max Planck Research School on Neuroscience of Communication: Function,

Structure, and Plasticity (IMPRS NeuroCom) by the Max Planck Society and the Predoc Award Program of the University of Leipzig.

ACKNOWLEDGMENTS

We thank all participants and the team at the LIFE study center, who made this study possible.

REFERENCES

- 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
- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edn.* Arlington, VA: American Psychiatric Publishing, USA.
- Amt für Statistik und Wahlen Leipzig (2020). Available online at: https://statistik.leipzig.de/statcity/agepyr_statsvg.aspx (accessed November 27, 2020).
- Baglio, F., Castelli, I., Alberoni, M., Blasi, V., Griffanti, L., Falini, A., et al. (2012). Theory of mind in amnesic mild cognitive impairment: an fMRI study. *J. Alzheimers Dis.* 29, 25–37. doi: 10.3233/JAD-2011-111256
- Bailey, P. E., Henry, J. D., and von Hippel, W. (2008). Empathy and social functioning in late adulthood. *Aging Ment. Health* 12, 499–503. doi: 10.1080/13607860802224243
- Baker, C., 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
- Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: on the dynamics between growth and decline. *Develop. Psychol.* 23, 611–626. doi: 10.1037/0012-1649.23.5.611
- 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., 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: 10.1111/1469-7610.00715
- Bölte, S. (2005). *Reading the Mind in the Eyes Test für Erwachsene (dt. Fassung) von S. Baron-Cohen.* Frankfurt/M: J.W. Goethe-Universität, Germany.
- Bora, E., Vahip, S., Gonul, A. S., Akdeniz, F., Alkan, M., Ogut, M., et al. (2005). Evidence for theory of mind deficits in euthymic patients with bipolar disorder. *Acta Psychiatr. Scand.* 112, 110–116. doi: 10.1111/j.1600-0447.2005.00570.x
- 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
- Cabinio, M., Rossetto, F., Blasi, V., Savazzi, F., Castelli, I., Massaro, D., et al. (2015). Mind-reading ability and structural connectivity changes in aging. *Front. Psychol.* 6:1808. doi: 10.3389/fpsyg.2015.01808
- Castelli, I., Baglio, F., Blasi, V., Alberoni, M., Falini, A., Liverta-Sempio, O., et al. (2010). Effects of aging on mindreading ability through the eyes: an fMRI study. *Neuropsychologia* 48, 2586–2594. doi: 10.1016/j.neuropsychologia.2010.05.005
- 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
- El Haj, M., Raffard, S., and Gely-Nargeot, M. C. (2015). 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’s dementia and normal aging. *Am. J. Neuroradiol.* 8, 421–426. doi: 10.2214/ajr.149.2.351
- Fischer, A. L., O’Rourke, N., and Thornton, W. L. (2016). Age differences in cognitive and affective theory of mind: concurrent contributions of neurocognitive performance, sex, and pulse pressure. *J. Gerontol. Series B Psychol. Sci. Soc. Sci.* 72, 71–81. doi: 10.1093/geronb/gbw088
- Gregory, C., Lough, S., Stone, V., Erzincliglu, 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
- Happé, F., Cook, J. L., and Bird, G. (2017). The structure of social cognition: in(ter)dependence of sociocognitive processes. *Annu. Rev. Psychol.* 68, 243–267. doi: 10.1146/annurev-psych-010416-044046
- 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
- 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
- Kirkland, R. A., Peterson, E., Baker, C. A., Miller, S., and Pulos, S. (2013). Metaanalysis reveals adult female superiority in “Reading the mind in the eyes” test. *North Am. J. Psychol.* 15, 449–458.
- Kung, K. (2020). Autistic traits, systemising, empathising, and theory of mind in transgender and non-binary adults. *Mol. Autism* 11:73. doi: 10.1186/s13229-020-00378-7
- 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 & memory. *J. Cereb. Blood Flow Metab.* 38, 996–1009. doi: 10.1177/0271678X17719380
- Kynast, J., Quinque, E. M., Polyakova, M., Luck, T., Riedel-Heller, S. G., Baron-Cohen, S., et al. (2020). Mindreading from the eyes declines with aging - evidence from 1,603 subjects. *Front. Aging Neurosci.* 12:550416. doi: 10.3389/fnagi.2020.550416
- 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
- 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 10000 deeply phenotyped adults in Germany. *BMC Public Health* 15:691. doi: 10.1186/s12889-015-1983-z
- Moosbrugger, H., and Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion, 2nd Edn.* Berlin/Heidelberg: Springer, Germany.
- Oakley, B., Brewer, R., Bird, G., and Catmur, C. (2016). Theory of mind is not theory of emotion: a cautionary note on the reading the mind in the eyes test. *J. Abnormal Psychol.* 125, 818–823. doi: 10.1037/abn0000182
- Olderbak, S., Wilhelm, O., Orlau, G., Geiger, M., Brennehan, M. W., and Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. *Front. Psychol.* 6:1503. doi: 10.3389/fpsyg.2015.01503
- 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. (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
- 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
- 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., Ettrich, B., Menz, M., and Zysset, S. (2010). Traumatic brain injury affects the frontomedian cortex – an event-related fMRI study on evaluative judgments. *Neuropsychologia* 48, 185–193. doi: 10.1016/j.neuropsychologia.2009.09.004
- 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 behavioural 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
- Statistisches Bundesamt Wiesbaden (2018). Available online at: <https://www-genesis.destatis.de> (accessed September 24, 2018).
- Vallat-Azouvi, C., Azouvi, P., Le-Bornec, G., and Brunet-Gouet, E. (2018). Treatment of social cognition impairments in patients with traumatic brain injury: a critical review. *Brain Injury* 22, 1–7. doi: 10.1080/02699052.2018.1531309

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Kynast, Polyakova, Quinque, Hinz, Villringer and Schroeter. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.