Measuring cognitive and affective Theory of Mind with the Italian Yoni task: normative data and short versions

Sara Isernia[1] · Federica Rossetto[1] · Valeria Blasi[1] · Davide Massaro[2] · Ilaria Castelli[3] · Cristian Ricci[4] · Simone Shamay-Tsoory[5] · Antonella Marchetti[2] · Francesca Baglio[1]

Abstract

Theory of Mind (ToM), a high-order human social cognition skill, is relevant for mental health and wellbeing and frequently impaired in many clinical conditions. To detect ToM difficulties, validated and standardized measures need to be adopted in clinical practice. The Yoni task provides a comprehensive evaluation of first-order, second-order, affective and cognitive mentalistic inferences. However, the Yoni task remains a time-demanding tool without normative data. To fill this gap, we explored the convergent and discriminant validity, the inter-item reliability and provided normative data of the Italian version of the Yoni task (98 items), and developed two short versions (48 and 36 items). For the validation, internal consistency and convergent/discriminant validity have been considered. Then, the effects of demographical variables on 98 item Yoni were explored, and normative data were provided. Our data demonstrated high reliability and a good convergent validity in all the Yoni versions. Then, two composite scores were proposed to obtain: 1) a global ToM level and 2) the balance between affective and cognitive ToM components, whose alteration represents a signature of several clinical conditions. The future standardization of the short versions of Yoni will allow a comprehensive and quick examination of ToM abilities, suitable for both research and clinical settings.

Introduction

Social cognition refers to any process involved in the detection, elaboration, and interpretation of social information, that is, the ability to represent other's intentions, emotions, desires and beliefs, and appropriately respond to them (Happé & Frith, 2014). These capabilities are crucial for successful communication and social interaction, with significant implications in mental health, wellbeing, and quality of life in both typical and atypical development (Adolphs, 2009; Cotter et al., 2018; Henry et al., 2016). In fact, social cognitive impairments are among the earliest and predominant clinical symptoms of many acute and chronic neurocognitive disorders (Adenzato & Poletti, 2013; Bora et al., 2015; Henry et al., 2016).

The relevance of social cognitive functioning assessment has been formally notified in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders, where social cognition is included as one of the six cores of neurocognitive domains (APA, 2013). Accordingly, different social cognitive assessment measures are now available for clinical use as part of wider neurocognitive batteries. However, differently from the five traditional neurocognitive domains (learning and memory, executive function, complex attention, perceptual-motor function, and language), grounded on a long history of validation of neuropsychological assessment tools, studies on psychometric characteristics, and normative data of social cognition tests are relatively few. This is especially true for high-order social cognition abilities, such as Theory of Mind (ToM) or mentalizing, the ability to think to one's own, and others' mental states to comprehend and predict their behaviours (Baglio & Marchetti, 2016; Call &

✉ Federica Rossetto
frossetto@dongnocchi.it

1   IRCCS Fondazione Don Carlo Gnocchi ONLUS, Milan, Italy

2   Unità di Ricerca sulla Teoria della Mente, Università Cattolica del Sacro Cuore, Milan, Italy

3   Department of Human and Social Science, Università Degli Studi Di Bergamo, Bergamo, BG, Italy

4   Africa Unit for Transdisciplinary Health Research (AUTHeR), North-West University, Potchefstroom, South Africa

5   Department of Psychology, University of Haifa, Haifa, Israel

Tomasello, 2008; Premack & Woodruff, 1978; Wimmer & Perner, 1983). The nature of this capability is complex and multi-component, embracing an affective and a cognitive component, referring respectively to emotions and intentions driving behaviour. Also, the meta-representation of mental states presents different levels of attribution, such as the first- and second-order inference (Kalbe et al., 2010; Shamay-Tsoory et al., 2005), which have been classically investigated with first and second-order false belief tasks (Baron-Cohen et al., 1985; Perner & Wimmer, 1985; Wimmer & Perner, 1983).

Although a substantial number of ToM tools are presented in the literature, most of them do not assess all these components of mentalizing. For example, the already mentioned False Belief task evaluates only the recursive level of mentalizing, without considering separately cognitive and affective components. On the other hand, the Faux Pas Recognition test (Roca et al., 2014; Stone et al., 1998), as well as the Story-based Empathy task (Dodich et al., 2015), assess both the affective and cognitive components without taking account recursive thinking (e.g., first and second order beliefs). Even the more ecological measures of ToM, based on multimedia stimuli, as the Toward the Assessment of Social Cognition test (TASIT, McDonald et al., 2004), and the Movie for the Assessment of Social Cognition (MASC, Dziobek et al., 2006) focus only on cognitive and affective ToM aspects, and not on the level of recursive inference. Moreover, to our knowledge, only the Faux Pas test and Story-based Empathy task have been provided with normative data (Delgado-Álvarez et al., 2021; Dodich et al., 2015). Furthermore, it has to be mentioned that nor the Faux Pas test nor the Story-based Empathy task has been validated. To date, the ToM test validated and mostly adapted for different languages and age groups remains a measure of affective mentalizing, the Eyes Test (Baron-Cohen et al., 2001; Fernandez-Abascal et al., 2013; Vellante et al., 2013).

Among the available ToM measures, the Yoni task (Shamay-Tsoory et al., 2007) is currently used as a complete computerized ToM measure to evaluate first and second-order, affective and cognitive mental state attributions. The Yoni task has been proved to be valuable in the research context. Different versions of the Yoni task have been applied in many clinical conditions: the 64-item version has been used to investigate cognitive and affective ToM in patients with localized brain lesions (Abu-Akel & Shamay-Tsoory, 2011), schizophrenia, and criminal offenders (Shamay-Tsoory et al., 2010); Bodden and colleagues (2010) adopted a 60-item version as a measure of mentalizing ability in patients with Parkinson Disease and healthy subjects. Finally, in the Italian context, Rossetto and colleagues (2018) assessed ToM abilities with the translation of the Yoni task, in its 98-item version, in subjects with Mild Cognitive Impairment and Parkinson's disease.

Considering the benefits of the Yoni task compared to other social cognitive tests, the Yoni task presents many advantages. Firstly, differently from the majority of social cognition tools, it allows a global comprehensive assessment of ToM, evaluating both cognitive and affective mental states, and also first and second-order beliefs. Also, Yoni presents simple visual stimuli with an ease of presentation modality purposively designed to minimize the influence of language, memory, and executive function on the subject's performance. This aspect is crucial considering that other social cognition measures have been reported to be influenced by the level in other neurocognitive domains (Lugnegård et al., 2013; McDonald et al., 2015; Schneideret al., 2012a, 2012b). Finally, divergently from story-based tests such as the Strange Stories (Happé, 1994), requiring a contextual interpretation of interactions, the simple stimuli of Yoni allow excluding the effect of social norms and autobiographical memory on the test score (Zalla & Korman, 2018).

Although the advantages related to Yoni task utilization, some issues need to be reported. Heterogeneous versions of the task are available in the literature, and neither validity (e.g., convergent and discriminant validity) and reliability (e.g., inter-item reliability) have been explored nor normative data have been presented. Moreover, to include the Yoni task into neuropsychological screening, the existing versions of the tool are not agile in their current form for the huge time-related demand.

According to these premises, the present study aimed at: a) exploring the convergent and discriminant validity, and inter-item reliability of the Italian version of the Yoni task score (Rossetto et al., 2018); b) providing normative data for the 98-item version in the Italian population; c) developing two short versions of the Yoni task from the 98-items version (Rossetto et al., 2018), balanced for its subdomains (first- and second-order, affective and cognitive components).

Methods

A secondary analysis was performed on data collected during previous researches (Isernia et al., 2019, 2020; Rossetto et al., 2018) and ongoing research. The IRCSS Don Gnocchi Foundation Ethic Committee and/or Catholic University of Sacred Heart Ethic Committee approved all the original research protocols, and data were collected in line with the Declaration of Helsinki. All participants read and signed the written informed consent before taking part in the study.

Participants

Participant's data were considered for the secondary analysis whether in line with the following inclusion criteria:

i) age > 18; ii) absence of neurological and/or psychiatric conditions assessed by the clinical interview; iii) absence of oncological conditions; iv) a global cognitive function as assessed by MoCA or MMSE within the normal range; v) absence of auditory and/or visual disability; vi) absence of pharmacological treatment affecting cognitive functions; vii) if under pharmacological therapy, a stable treatment for at least three months before the participation in the research; viii) a high performance in the control items of ToM tasks.

All these subjects were recruited at IRCCS Don Gnocchi Foundation of Milan and Catholic University of Sacred Heart of Milan. They were students, staff of the clinic, and volunteers that were contacted by mail and/or telephone by a researcher.

ToM assessment

The Italian adapted version of the *Yoni Task* (Rossetto et al., 2018) was used to test ToM abilities. This task consists of 98 items, 84 mental items, and 14 physical (control) items showing a face named "Yoni" and four colored pictures placed around the face referring to different semantic categories (e.g., fruit, animals, means of transport) or faces. The items are presented on a computer screen by a software (e.g., e-PRIME) that allows to register time response as well as accuracy. The task requires the participant to choose the correct item to which Yoni refers based on a sentence on the top of the screen and specific cues such as Yoni's eye gaze and/or facial expression, and/or the eye gaze/facial expression of faces around him. Participants were instructed to choose the correct answer among four alternatives by selecting them pointing on it with the computer mouse as fast as possible. All items differ in the complexity of the meta-representation they require, i.e., first-order levels ($N = 24$ items) or second-order levels ($N = 60$), and in the evaluation of affective ToM ("Yoni likes...", $N = 48$ items) or cognitive ToM ("Yoni is thinking of...", $N = 36$ items) (Fig. 1). The performance was rated separately for accuracy and response times. For the ToM accuracy score (ACC), only mental items were considered; each item scored 1 if correct and 0 if wrong, for a total score of 84. Also, the ACC scores gained from each subcategory were summed to obtain four subtotals: the total of first-order items (1ORD, range 0–24), the total of second-order items (2ORD, range 0–60), the total of affective items (AFF, range 0–48), and the total of cognitive items (COG, range 0–36). Response time (seconds elapsing between the stimulus presentation and the subject answer), total score (RT) was obtained by averaging only the response time of correct items. As ACC scores, separate scores related to subcategories of the scale were computed for RT.

The Italian version of the *Reading the Mind in the Eyes* test *(*RMET, Baron-Cohen et al., 2001; Vellante et al., 2013) was considered a measure of advanced ToM based on judging mental states from the eyes. The RMET comprises 36 items, presenting black-and-white photographs of the eye region of males and females displaying a certain basic or complex emotion. Each stimulus is reported with four forced-choice alternatives consisting of four states of mind. The subject is invited to choose the answer that best fits with the state of mind depicted in the photograph. Each item scores 0–1, for a total score that ranges 0–36. A gender recognition task (Gender Test, GT) is also performed with the same items as a control task to exclude visual disability affecting task performance. For the present study, a computerized version of the task has been administered.

Statistical analysis

Jamovi 1.2.27 and IBM Statistical Package for Social Science v. 26 software were used for the statistical analysis.

Description of the participants characteristics Summary statistics, such as means, and standard deviations, were run to report the demographic characteristics of participants. For each variable included in the analyses, normality distribution and modality were considered, and consequently, parametrical or non-parametrical analyses were run.

Validity and Reliability of 98-item Yoni Cronbach's alpha and McDonald's omega were computed to verify the internal consistency of Yoni. Inter-item reliability was also tested by split-half parallel reliability Spearman-Brown $\varrho_{SP.}$ Convergent and discriminant validity was tested by performing correlations (Spearman's ρ) between Yoni and RMET, and Yoni and Gender Test.

Standardization of the 98-item Yoni task and normative data Simultaneous multivariate regressions were run to explore the role of demographic variables (sex, age, and years of education) as possible predictors of Yoni accuracy and response time. Then, the contribution of predictors in the regression model was included in a formula to adjust raw scores to exclude the effects of demographic variables on Yoni accuracy and response time. Then, the specific adjustment to be added or subtracted from the raw score based on the subject's sex, age, and years of education was computed to provide adjustment score tables. No adjustment was suggested for the upper and inner limit of Yoni's accuracy raw score. Assumptions for the application of linear regression were graphically checked by means of residual distributions and homoscedasticity (histogram plot of residuals, plots of residuals vs fitted, Normal Q-Q plot, square root of standardized residuals vs. fitted values, standardized residuals vs. leverage) and appropriate statistical tests (Breush-Pagan Test and Non-constant Variance Score Test). All regression models were applied to inverse rank Blom's transformed data to improve the normality of variable distribution and
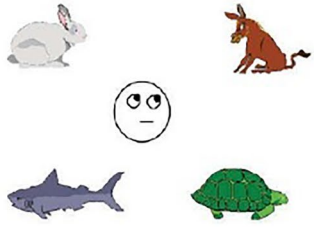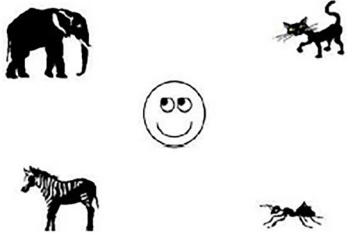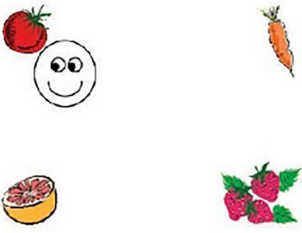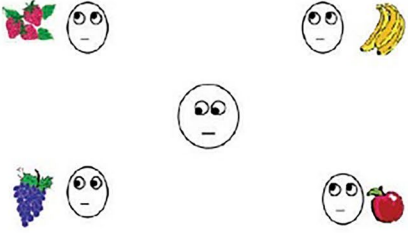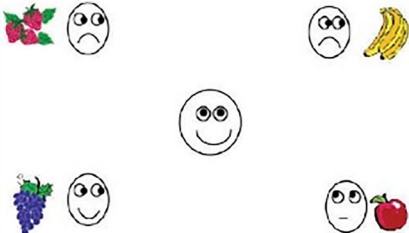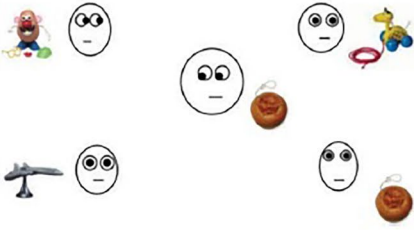
Fig. 1 Sample of items from the Yoni task: first- and second-order, cognitive and affective mental inference and physical (control) items (Source: from Rossetto et al., 2018)

the adequacy of linear model approach. Finally, an adjusted total score of Yoni was computed separately for accuracy and response time, considering only mental and not physical items. For both accuracy and response time, the total score was set to range 0–1. Accuracy total score was computed by summing adjusted scores of first-order and second-order items and dividing the sum for the total number of mental items (ACC_TOT = *first-order adjusted ToM + second-order adjusted ToM / N_i*). The total response time score was created by averaging the sum of the first and second-order adjusted ToM response time on the total available time minus the minimum time necessary for the performance (RT_ToT = *1-((((First-order adjusted ToM RT—RT_{min}) / RT_i) + ((Second-order adjusted ToM RT—RT_{min}) / RT_i))) / 2).*

Furthermore, an additional score was proposed to detect the balance between the level of affective and cognitive ToM, separately for accuracy and response time, as follows:

Cognitive/Affective Accuracy index (CA$_A$) = (Affective Adjusted ToM / N_i)

−(Cognitive Adjusted ToM / N_i)/

(Affective Adjusted ToM / N_i)

+(Cognitive Adjusted ToM / N_i);

Cognitive/Affective Response Time index (CA$_{RT}$) = (((Affective Adjusted ToM / RT_i)

−((Cognitive Adjusted ToM /RT_i)) ∗ −1).

A CA$_A$/CA$_{RT}$ index near 0 suggested a balance between affective and cognitive ToM; a positive index indicated a

higher affective ToM level than cognitive ToM, and a negative index suggested a higher level of cognitive ToM than affective ToM.

Summary statistics were performed to show Italian normative data of the Yoni 98 item version.

Development of Italian Yoni short forms Cronbach's alpha, McDonald's omega, Cronbach's alpha if dropped and McDonald's omega if dropped were considered for all 98 item of Yoni to select the items' pool for the development of the Yoni short versions. To this purpose, a balanced number of items for each ToM component (first-order; second-order; affective; cognitive) was considered, and only items contributing to a high Cronbach's alpha, McDonald's omega of the scale (alpha/omega 0.8–0.9) were included in the short version.

Sample size calculation To compute sample size adequacy, statistical power for correlation coefficients and Cronbach's alpha was performed considering statistical tests aimed to detect correlation and Cronbach's alpha coefficients of 0.75 with a null coefficient of 0.6. Notably, we considered power calculations for correlations as valid also for the Cronbach's alpha coefficients as they distribute as the test–retest correlation of the scale scores (Heo et al., 2015). Power calculations for multivariate regression were performed considering medium Cohen's f^2 effect size. According to the above power calculations, the sample size used ($N = 175$) is sufficient to guarantee a statistical power ($1-\beta$) above 90% for all the statistical tests here reported.

Results

Participants

Data of 175 healthy subjects (65 males, mean age $= 38.4 \pm 20.6$, mean years of education $= 14.8 \pm 3.17$) were included in the analysis. Table 1 shows the characteristics of the subjects for age groups.

Table 1 Characteristics of the participants for age groups

	N (Male/Female)	Age (M,sd)	Education (y) (M,sd)
Age group			
18–29	99 (32/67)	22.5,3.14	15.3,2.48
30–49	15 (8/7)	37.1,7.39	16.7,2.87
50–59	24 (12/12)	55.1, 2.86	15.0, 2.87
60–69	16 (5/11)	64.7, 2.85	14.0, 3.61
70+	21 (8/13)	75.4, 3.81	11.0, 3.46

M mean; *sd* standard deviation; *y* years

Internal consistency, inter-item reliability, convergent and discriminant validity of Yoni-98

The 98-item Yoni version showed a mean item accuracy of 0.90 ± 0.12, a high Cronbach's α of 0.90, and a high McDonald's $\omega = 0.89$. The tool demonstrated a high inter-item reliability: split-half parallel reliability Spearman-Brown ϱ_{SP} median $= 0.90$, 95% HDI $= 0.85$–0.93. Also, it showed a good convergent validity by reporting a statistically significant correlation with RMET (Yoni total: $\varrho = 0.260$, $p < 0.001$; 2ORD: $\varrho = 0.252$, $p < 0.001$; AFF: $\varrho = 0.287$, $p < 0.001$; COG: $\varrho = 0.197$, $p = 0.009$; 1ORD COG: $\varrho = 0.170$, $p = 0.025$; 2ORD AFF: $\varrho = 0.286$, $p < 0.001$; 2ORD COG: $\varrho = 0.187$, $p = 0.013$). Finally, a good discriminant validity was reported with a no significant correlation with GT (Yoni total: $\varrho = 0.04$, $p = 0.872$; 1ORD: $\varrho = 0.02$, $p = 0.797$, 2ORD: $\varrho = 0.01$, $p = 0.797$; AFF: $\varrho = 0.08$, $p = 0.312$; COG: $\varrho = -0.05$, $p = 0.512$; 1ORD AFF: $\varrho = 0.02$, $p = 0.797$; 2ORD AFF: $\varrho = 0.07$, $p = 0.361$; 1ORD COG: $\varrho = 0.04$, $p = 0.585$; 2ORD COG: $\varrho = -0.04$, $p = 0.584$).

Effects of sex, age and years of education on Yoni-98

Table 2 shows the simultaneous regression models (Table 2).

The regression models revealed age as a significant predictor on second-order, affective ToM accuracy Yoni scores and all Yoni response time scores. Years of education were demonstrated to be significant predictors on all Yoni-98 accuracy scores except for first-order ToM, but not on response time scores. No predictive effect of sex on Yoni scores was observed.

Normative values to convert raw scores are proposed in Table 3, and an example of a scoring sheet is provided in Table 4.

Adjusted Yoni scores in the Italian population

Table 5 reports summary statistics of adjusted total scores of Yoni-98.

Development of Yoni short versions

By selecting items that highly contributed to the internal consistency of the scale, two short versions of Yoni were created (see Table 6), a 48-item (Yoni-48) and a 36-item (Yoni-36) versions. Both the Yoni short versions presented a high internal consistency (Yoni-48: $\alpha = 0.90$/ $\omega = 0.90$; Yoni-36: $\alpha = 0.88$/ $\omega = 0.86$) and a good mean item accuracy (Yoni-48: M $= 0.87 \pm 0.14$; Yoni-36: M $= 0.89 \pm 0.14$). Also, both version revealed a high convergent validity, correlating with RMET (Yoni-48; total: $\varrho = 0.253$, $p < 0.001$; 2ORD: $\varrho = 0.251$, $p < 0.001$; AFF: $\varrho = 0.293$, $p < 0.001$; COG: $\varrho = 0.175$, $p = 0.021$; 2ORD AFF: $\varrho = 0.295$,

Table 2 Simultaneous regression models of demographic variables on 98-item Yoni

	Predictors	β	S.E	95% CI lower,upper	t	targeted p	omnibus p	R²
Accuracy								
1ORD	Sex	0.01	0.16	-0.33,0.32	0.05	0.963	0.327	0.020
	Age	-0.01	0.00	-0.02–0.00	-1.82	0.071		
	Education(y)	-0.02	0.03	-0.02,0.03	-0.80	0.425		
	Intercept	0.58	0.45	-0.42,1.59	1.15	0.250		
2ORD	Sex	-1.86	1.08	-0.26,3.98	-1.73	0.085	<0.001	0.164
	Age	-0.07	0.03	-0.12,-0.02	-2.87	0.005		
	Education(y)	0.63	0.17	0.29,0.97	3.64	<0.001		
	Intercept	46.19	2.97	37.74,50.94	15.57	<0.001		
AFF	Sex	-0.18	0.15	-0.13,0.48	-1.16	0.249	<0.001	0.141
	Age	-0.01	0.00	-0.02,-0.00	-3.18	0.002		
	Education(y)	0.07	0.02	0.02,0.12	2.82	0.005		
	Intercept	-0.51	0.42	-1.63,0.25	-1.21	0.229		
COG	Sex	-0.25	0.16	-0.06,0.56	-1.60	0.112	0.001	0.088
	Age	-0.01	0.00	-0.01,0.00	-1.57	0.118		
	Education(y)	0.07	0.02	0.02,0.12	2.84	0.005		
	Intercept	-0.74	0.43	-1.96,-0.02	-1.70	0.091		
Response Time								
1ORD	Sex	-0.20	0.14	-0.08,0.49	-1.39	0.165	<0.001	0.223
	Age	0.02	0.00	0.01,0.03	6.38	<0.001		
	Education(y)	-0.01	0.02	-0.05,0.04	-0.33	0.743		
	Intercept	-0.66	0.40	-1.76,0.03	-1.65	0.101		
2ORD	Sex	-0.04	0.13	-0.22,0.31	-0.34	0.731	<0.001	0.352
	Age	0.03	0.00	0.02,0.03	8.64	<0.001		
	Education(y)	-0.02	0.02	-0.06,0.02	-1.01	0.315		
	Intercept	-0.72	0.37	-1.58,0.05	-1.96	0.051		
AFF	Sex	-0.12	0.14	-0.15,0.38	-0.86	0.389	<0.001	0.328
	Age	0.03	0.00	0.02,0.03	8.25	<0.001		
	Education(y)	-0.02	0.02	-0.06,0.02	-0.81	0.418		
	Intercept	-0.72	0.37	-1.67,-0.01	-1.93	0.056		
COG	Sex	-0.06	0.13	-0.20,0.32	-0.45	0.650	<0.001	0.345
	Age	0.03	0.00	0.02,0.03	8.56	<0.001		
	Education(y)	-0.02	0.02	-0.06,0.02	-0.92	0.361		
	Intercept	-0.74	0.37	-1.62,0.02	-2.00	0.048		

1ORD First-Order ToM; *2ORD* Second-Order ToM; *AFF* Affective ToM; *COG* Cognitive ToM; *S.E.* Standard Error

$p < 0.001$; 2ORD COG: ϱ = 0.185, $p = 0.014$. Yoni-36; total: ϱ = 0.233, $p = 0.002$; 2ORD: ϱ = 0.229, $p = 0.002$; AFF: ϱ = 0.286, $p < 0.001$; COG: ϱ = 0.151, $p = 0.046$; 2ORD AFF = ϱ = 0.287, $p < 0.001$; 2ORD COG: ϱ = 0.155, $p = 0.041$), and a good discriminant validity, showing no correlation with GT (Yoni-48; total: ϱ = -0.020, $p = 0.789$; 1ORD: ϱ = 0.062, $p = 0.414$; 2ORD: ϱ = -0.028, $p = 0.712$; AFF: ϱ = 0.051, $p = 0.503$; COG: ϱ = -0.033, $p = 0.662$; 1ORD AFF: ϱ = 0.061, $p = 0.424$; 1ORD COG: ϱ = 0.071, $p = 0.351$; 2ORD AFF: ϱ = 0.032, $p = 0.674$; 2ORD COG: ϱ = -0.023, $p = 0.762$. Yoni-36; total: ϱ = -0.020, $p = 0.772$; 2ORD: ϱ = -0.039, $p = 0.612$; AFF: ϱ = 0.041, $p = 0.586$;

COG: ϱ = -0.040, $p = 0.601$; 2ORD AFF: ϱ = 0.002, $p = 0.979$; 2ORD COG: ϱ = -0.025, $p = 0.745$).

Table 7 depicts the number of items dedicated to each ToM component evaluated in the full and short versions.

Discussion

Yoni constitutes a useful ToM measure to investigate neurocognitive impairment in atypical conditions. However, the absence of validation and normative data makes it less suitable for adoption in clinical practice. To this purpose,

Table 3 Adjustment for age and Years of education of Yoni-98 raw scores

			Age									
			18–29		30–49		50–59		60–69		70–80	
			female	male	female	male	female	male	female	male	female	male
		Education										
Accuracy	1ORD	5	0.118	0.111	0.216	0.209	0.321	0.314	0.391	0.384	0.465	0.458
		8	0.055	0.048	0.153	0.146	0.258	0.251	0.328	0.321	0.402	0.395
		13	-0.050	-0.057	0.048	0.041	0.153	0.146	0.223	0.216	0.297	0.290
		16	-0.113	-0.120	-0.015	-0.022	0.090	0.083	0.160	0.153	0.234	0.227
	2ORD	5	4.524	6.386	5.546	7.408	6.641	8.503	7.371	9.233	8.137	9.999
		8	2.640	4.502	3.662	5.524	4.757	6.619	5.487	7.349	6.253	8.115
		13	-0.500	1.362	0.522	2.384	1.617	3.479	2.347	4.209	3.113	4.975
		16	-2.384	-0.522	-1.362	0.500	-0.267	1.595	0.463	2.325	1.229	3.091
	AFF	5	0.456	0.633	0.624	0.801	0.804	0.981	0.924	1.101	1.050	1.227
		8	0.249	0.426	0.417	0.594	0.597	0.774	0.717	0.894	0.843	1.020
		13	-0.096	0.081	0.072	0.249	0.252	0.429	0.372	0.549	0.498	0.675
		16	-0.303	-0.126	-0.135	0.042	0.045	0.222	0.165	0.342	0.291	0.468
	COG	5	0.535	0.633	0.619	0.801	0.709	0.981	0.769	1.101	0.832	1.227
		8	0.319	0.426	0.403	0.594	0.493	0.774	0.553	0.894	0.616	1.020
		13	-0.041	0.081	0.043	0.249	0.133	0.429	0.193	0.549	0.256	0.675
		16	-0.257	-0.126	-0.173	0.042	-0.083	0.222	-0.023	0.342	0.040	0.468
Response Time	1ORD	5	0.281	0.078	-0.027	-0.230	-0.357	-0.560	-0.577	-0.780	-0.808	-1.011
		8	0.305	0.102	-0.003	-0.206	-0.333	-0.536	-0.553	-0.756	-0.784	-0.987
		13	0.345	0.142	0.037	-0.166	-0.293	-0.496	-0.513	-0.716	-0.744	-0.947
		16	0.369	0.166	0.061	-0.142	-0.269	-0.472	-0.489	-0.692	-0.720	-0.923
	2ORD	5	0.160	0.114	-0.218	-0.264	-0.623	-0.669	-0.893	-0.939	-1.177	-1.223
		8	0.223	0.177	-0.155	-0.201	-0.560	-0.606	-0.830	-0.876	-1.114	-1.160
		13	0.328	0.282	-0.050	-0.096	-0.455	-0.501	-0.725	-0.771	-1.009	-1.055
		16	0.391	0.345	0.013	-0.033	-0.392	-0.438	-0.662	-0.708	-0.946	-0.992
	AFF	5	0.215	0.098	-0.163	-0.280	-0.568	-0.685	-0.838	-0.955	-1.121	-1.238
		8	0.269	0.152	-0.109	-0.226	-0.514	-0.631	-0.784	-0.901	-1.067	-1.184
		13	0.359	0.242	-0.019	-0.136	-0.424	-0.541	-0.694	-0.811	-0.977	-1.094
		16	0.413	0.296	0.035	-0.082	-0.370	-0.487	-0.640	-0.757	-0.923	-1.040
	COG	5	0.175	0.114	-0.203	-0.264	-0.608	-0.669	-0.878	-0.939	-1.162	-1.223
		8	0.235	0.174	-0.143	-0.204	-0.548	-0.609	-0.818	-0.879	-1.102	-1.163
		13	0.335	0.274	-0.043	-0.104	-0.448	-0.509	-0.718	-0.779	-1.002	-1.063
		16	0.395	0.334	0.017	-0.044	-0.388	-0.449	-0.658	-0.719	-0.942	-1.003

For accuracy score: for 2ORD, a raw score of 0 or 60 has not to be adjusted; for AFF, a raw score of 0 or 48 has not to be adjusted; for COG, a raw score of 0 or 36 has not to be adjusted

1ORD First-Order ToM; *2ORD* Second-Order ToM; *AFF* Affective ToM; *COG* Cognitive ToM

our first aim consisted in exploring the psychometric characteristic of the Italian version of the Yoni task (Yoni-98; Rossetto et al., 2018) in terms of both validity and reliability.

As expected, our results revealed the Yoni-98 score as highly valid and reliable, reporting an adequate internal consistency for both clinical and research purposes (Streiner, 2003; Tavakol & Dennick, 2011) and a significant association with the RMET, which is one of the most used advanced ToM tests (Vellante et al., 2013).

Then, to enhance the reliable application of the tool, normative data were generated, and two composite scores were proposed to detect: 1) the global level of ToM and 2) the balance between affective and cognitive ToM ability.

To our knowledge, only a few works on ToM measures reported validation (e.g., Olderbak et al., 2015; Vellante et al., 2013) as well as normative data (e.g., Delgado-Álvarez et al., 2021; Dodich et al., 2015). This is the first study reporting normative data on the Italian version of Yoni, with

Table 4 Yoni-98 scoring sheet example

Accuracy	RAW	ADJUSTED	Response Time	RAW	ADJUSTED
ACC_1ORD (0–24)			RT_1ORD		
ACC_2ORD (0–60)			RT_2ORD		
ACC_AFF (0–48)			RT_AFF		
ACC_COG (0–36)			RT_COG		
ACC_PHYSICAL (0–14)			ACC_PHYSICAL		
ACC_TOT (0–1)			RT_TOT (0–1)		
CA_A			CA_{RT}		

CA_A cognitive/affective accuracy balance score; CA_{RT} cognitive/affective ToM response time balance; *ACC_1ORD* first-order ToM accuracy; *ACC_2ORD* second-order ToM accuracy; *ACC_AFF* affective ToM accuracy; *ACC_COG* cognitive ToM accuracy; *ACC_TOT* Total ToM accuracy; *RT_1ORD* first-order ToM response time; *RT_2ORD* second-order ToM response time; *RT_AFF* affective ToM response time; *RT_COG* cognitive ToM response time; *RT_TOT* Total ToM response time

Formulas: $ACC\_TOT = (1ORD_{adj} + 2ORD_{adj})/84$; $CA_A = ((AFF_{adj}/48) - (COG_{adj}/36) / (AFF_{adj}/48) + (COG_{adj}/36))$; RT_TOT = *1-((((First-order adjusted ToM RT—RT_{min}) / RT_i) + ((Second-order adjusted ToM RT—RT_{min}) / RT_i))) / 2)*; CA_{RT} = *(((Affective Adjusted ToM / RT_i)—((Cognitive Adjusted ToM / RT_i))*-1)*

Table 5 Italian normative data of Yoni task

	YONI-98	
	ACC	RT(*s*)
Adjusted scores		
Mean, sd	0.89,0.09	0.79,0.10
Median	0.92	0.81
25th,75th percentile	0.85,0.96	0.73,0.85
CA		
Mean, sd	0.81,0.19	0.04,0.05
Median	0.87	0.03
25th,75th percentile	0.73, 0.95	0.01,0.06

ACC accuracy; *ADJ* adjusted; *CA* cognitive/affective indexes; *M* mean; *RT(s)* Response Time (seconds); *sd* standard deviation; *Yoni-98* 98-items version of Yoni

many implications in terms of theoretical values and practical contributions.

Especially, disposing of normative data allows exploring the nature of human social behaviour processes net of confounding variables such as demographical characteristics. In fact, different contributions attest to the link between demographical variables and social cognition performance (Baksh et al., 2018; Chiasson et al., 2017; Isernia et al., 2020; Rosi et al., 2016), highlighting the need to consider this association in ToM test scores.

Finally, our work provides formulas to compute the total scores of the Yoni task, both in terms of accuracy and response time. The ratio underlying the creation of these scores responds to the complex nature of ToM processes. In fact, different orders of ToM meta-representations, such as the first and second-order ToM inference, are recognized as

different levels of mentalizing competence, with a specific role in ToM achievement during child development (Valle et al., 2015; Wimmer & Perner, 1983). Our data revealed no effects of sex, years of education and age on the accuracy of the first-order ToM items. The absence of an influence of demographical variables on first-order mentalizing is in line with developmental research (Flavell, 1999; Hughes & Leekam, 2004; Repacholi & Slaughter, 2004), demonstrating the achievement of this ability already in the pre-school age. Instead, as expected, we found that only age was a predictor of the response time of first-order ToM items, mirroring the age-related behavioural slowing in sensorimotor tasks (Yordanova et al., 2004). Concerning the second-order ToM inferences, we highlighted the predictive role of age and years of education in both accuracy and response time. This evidence supports the complexity of this ToM component, resorting to high cognitive load processes, with a predominance of executive functions (Sandoz et al., 2014). Based on these findings, we generated normative data separately for first, and second-order ToM inference and subsequentially computed the total score.

Moreover, in line with the neuropsychological and neuroscience evidence demonstrating the co-existence of affective and cognitive ToM, with related separate brain networks and neurochemical systems (Abu-Akel & Shamay-Tsoory, 2011), new indexes have been proposed. In details, the CA_A and CA_{RT} index have been conceived to detect the balance between these two components.

The practical contributions of this work are vast. Especially, normative data are essential to include the Italian Yoni task into a battery for the neurocognitive abilities assessment. These scores might be relevant to detect a dissociate impairment in ToM function in the clinical

Table 6 Item pool of Yoni-48 and Yoni-36

	Yoni-98 item	ToM Component	α if dropped	ω if dropped	Yoni-48	Yoni-36
1ORD α=0.718 ω=0.750	3	AFF	0.724	0.753	x	x
	5	COG	0.713	0.749	x	x
	8	COG	0.707	0.746	x	x
	9	AFF	0.706	0.723	x	x
	10	COG	0.674	0.707	x	x
	12	AFF	0.718	0.753	x	x
	13	AFF	0.695	0.722	x	x
	20	AFF	0.704	0.728	x	x
	21	COG	0.701	0.715	x	x
	22	COG	0.716	0.756	x	x
	25	AFF	0.711	0.744	x	x
	30	AFF	0.688	0.741	x	x
	32	AFF	0.676	0.726	x	x
	36	COG	0.708	0.752	x	x
	54	COG	0.716	0.750	x	x
	55	COG	0.699	0.740	x	x
2ORD α=0.903 ω=0.906	29	AFF	0.899	0.901	x	x
	33	AFF	0.899	0.901	x	x
	34	AFF	0.900	0.901	x	x
	37	AFF	0.901	0.902	x	x
	38	AFF	0.901	0.903	x	x
	44	AFF	0.903	0.904	x	
	46	AFF	0.902	0.904	x	
	53	AFF	0.901	0.903	x	x
	61	AFF	0.902	0.903	x	
	63	COG	0.901	0.903	x	
	64	COG	0.901	0.903	x	
	68	COG	0.901	0.903	x	
	71	COG	0.896	0.899	x	x
	72	COG	0.898	0.901	x	x
	73	AFF	0.902	0.904	x	
	74	COG	0.903	0.905	x	
	75	AFF	0.899	0.901	x	x
	76	AFF	0.903	0.906	x	
	79	COG	0.897	0.899	x	x
	80	COG	0.900	0.902	x	x
	86	COG	0.902	0.904	x	
	87	COG	0.900	0.902	x	x
	91	AFF	0.899	0.901	x	x
	94	COG	0.896	0.899	x	x
	95	COG	0.896	0.899	x	x
	98	COG	0.898	0.900	x	x
					α=0.90 ω=0.90	α=0.88 ω=0.86

1ORD First-Order ToM; *2ORD* Second-Order ToM; *AFF* Affective ToM; *COG* Cognitive ToM; *Yoni-36* 36-items short version of Yoni; *Yoni-48* 48-items short version of Yoni; *Yoni-98* 98-items version of Yoni

For the Yoni-48 the following physical items have been selected: 6,7,15,60,77,93. For the Yoni-36 the following physical items have been selected: 6,7, 60, 93

Table 7 Structure of the Yoni versions

	Accuracy (M, sd)	Cronbach's alpha	McDonald's ω	1ORD (N)		2ORD (N)		PHYSICAL (N)	
				AFF	COG	AFF	COG	1ORD	2ORD
Yoni-98	0.90,0.12	0.90	0.89	12	12	36	24	8	6
Yoni-48	0.87,0.14	0.90	0.90	8	8	13	13	3	3
Yoni-36	0.89,0.14	0.88	0.86	8	8	8	8	2	2

1ORD First-Order ToM; *2ORD* Second-Order ToM; *AFF* Affective ToM; *COG* Cognitive ToM; *M* mean; *N* number; *sd* standard deviation; *Yoni-98* 98-items version of Yoni; *Yoni-48* 48-items short version of Yoni; *Yoni-36* 36-items short version of Yoni

setting. Indeed, several contributions attested the dissociation between affective and cognitive components of ToM in atypical development conditions (e.g., Isernia et al., 2019, 2020; Poletti et al., 2011; Rossetto et al., 2018): while some neurodegenerative diseases, such as Alzheimer and Parkinson's disease, show a specific impairment in cognitive but not affective ToM (Kemp et al., 2012; Rossetto et al., 2018), psychiatric disorders, as schizophrenia, present the opposite trend, with heavier damage in the affective mentalizing (Shamay-Tsoory et al., 2007). Instead, as reported by our data, a dissociation between the two ToM components does not emerge in healthy populations, reporting a CA index near 0, suggesting the balance between affective and cognitive mentalizing in typical lifespan conditions.

Finally, to answer the need to provide quick tools for agile use during the neuropsychological assessment, we developed and validated two short versions of the Yoni task from the 98-items version, Yoni-48, and Yoni-36. These short versions have been created also with the purpose to balance the number of items assessing separate ToM subdomains, such as first- and second-order, affective, and cognitive components. Focusing on the psychometric proprieties of these two versions, we observed high reliability and a good convergent validity in both Yoni-48 and Yoni-36. In particular, while Yoni-48 showed higher reliability ($\alpha = 0.90$), suitable for the adoption of the tool in clinical practice, Yoni-36 offers an optimal balance between ToM components assessed (8 items for each ToM domain). Both versions are able to provide a fast and complete evaluation of mentalistic abilities, essential features for the adoption in the clinic of these instruments.

Future contributions should provide evidence on the construct validity of the Yoni task by performing factor analysis. Also, normative data for the Yoni short versions both in the Italian healthy and clinical population have to be provided to adopt these instruments for the assessment of ToM in populations presenting social cognition deficit, such as people with schizophrenia, autism spectrum disorder, and neurodegenerative conditions. Also, the sensitivity and specificity of Yoni full and short versions should be investigated to reveal the potential diagnostic accuracy of the tool. Finally, the development of parallel forms of the Yoni tasks would offer the possibility to longitudinally evaluate ToM abilities in clinical practice, also in terms of rehabilitation outcomes.

This contribution is not without limitations: future works are needed to replicate these findings with a larger population, characterized with additional demographic factors, such as occupational background, social participation in daily life, etc. Also, future studies will need to test the convergent validity with additional ToM tools, assessing different ToM components, such as both affective and cognitive mentalizing, to confirm the results of the present work.

In conclusion, for the first time, this work presented a validation of the Italian version of the Yoni task and two short-versions of the test by proposing normative data and composite score formula. The implications of this contribution are vast both for research purposes and clinical practice. In fact, short versions with appropriate psychometric properties will allow a comprehensive examination of ToM abilities in a setting of neuropsychological screening where time is short, helping to plan and monitor a psychosocial rehabilitation intervention adequately (Rossetto et al., 2020).

Declarations

Conflict of interest No potential conflict of interest was reported by the authors.

References

American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM V)*. American Psychiatric.

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

Adenzato, M., & Poletti, M. (2013). Theory of Mind abilities in neurodegenerative diseases: An update and a call to introduce mentalizing tasks in standard neuropsychological assessments. *Clinical Neuropsychiaty, 10*(5), 226–234.

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology, 60*, 693–716.

Baglio, F., & Marchetti, A. (2016). Editorial: When (and How) Is Theory of Mind Useful? Evidence from Life-Span Research. *Front Psychol, 7*, 1425.

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

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

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.

Bodden, M. E., Mollenhauer, B., Trenkwalder, C., Cabanel, N., Eggert, K. M., Unger, M. M., et al. (2010). Affective and cognitive Theory of Mind in patients with parkinson's disease. *Parkinsonism & Related Disorders, 16*(7), 466–470.

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.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187–192.

Chiasson, V., Vera-Estay, E., Lalonde, G., Dooley, J. J., & Beauchamp, M. H. (2017). Assessing social cognition: Age-related changes in moral reasoning in childhood and adolescence. *The Clinical Neuropsychologist, 31*(3), 515–530.

Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C. Y., & Barnett, J. H. (2018). Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neuroscience and Biobehavioral Reviews, 84*, 92–99.

Delgado-Álvarez, A., Pytel, V., Delgado-Alonso, C., Olbrich-Guzmán, C. M., Cortés-Martínez, A., Moreno-Ramos, T., et al. (2021). Development, Spanish Normative Data, and Validation of a Social Cognition Battery in Prodromal Alzheimer's Disease and Multiple Sclerosis. *Archives of Clinical Neuropsychology, 36*(5), 711–722.

Dodich, A., Cerami, C., Canessa, N., Crespi, C., Iannaccone, S., Marcone, A., et al. (2015). A novel task assessing intention and emotion attribution: Italian standardization and normative data of the Story-based Empathy Task. *Neurological Sciences, 36*(10), 1907–1912.

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

Fernández-Abascal, E. G., Cabello, R., Fernández-Berrocal, P., & Baron-Cohen, S. (2013). Test-retest reliability of the 'Reading the Mind in the Eyes' test: a one-year follow-up study. *Molecular Autism, 4*(1), 33. https://doi.org/10.1186/2040-2392-4-33

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology, 50*, 21–45.

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped and normal children and adults. *Journal of Autism and Developmental Disorders, 24*, 129–154. https://doi.org/10.1007/BF02172093

Happé, F., & Frith, U. (2014). Annual research review: Towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry, 55*(6), 553–557.

Henry, J. D., von Hippel, W., Molenberghs, P., Lee, T., & Sachdev, P. S. (2016). Clinical assessment of social cognitive function in neurological disorders. *Nature Reviews. Neurology, 12*(1), 28–39.

Heo, M., Kim, N., & Faith, M. S. (2015). Statistical power as a function of Cronbach alpha of instrument questionnaire items. *BMC Medical Research Methodology, 15*, 86. https://doi.org/10.1186/s12874-015-0070-6

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? *Social Development, 13*(4), 590–619.

Isernia, S., Baglio, F., d'Arma, A., Groppo, E., Marchetti, A., & Massaro, D. (2019). Social Mind and Long-Lasting Disease: Focus on Affective and Cognitive Theory of Mind in Multiple Sclerosis. *Frontiers in Psychology, 10*, 218.

Isernia, S., Cabinio, M., Pirastru, A., Mendozzi, L., Di Dio, C., Marchetti, A., et al. (2020). Theory of mind network in multiple Sclerosis: A double disconnection mechanism. *Social Neuroscience, 15*(5), 544–557.

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

Kemp, J., Després, O., Sellal, F., & Dufour, A. (2012). Theory of Mind in normal ageing and neurodegenerative pathologies. *Ageing Research Reviews, 11*(2), 199–219.

Lugnegård, T., Unenge Hallerbäck, M., Hjärthag, F., & Gillberg, C. (2013). Social cognition impairments in Asperger syndrome and schizophrenia. *Schizophrenia Research, 143*(2–3), 277–284. https://doi.org/10.1016/j.schres.2012.12.001

McDonald, S., Fisher, A., Togher, L., Tate, R., Rushby, J., English, T., Kelly, M., Mathersul, D., Froreich, F., & Francis, H. (2015). Adolescent Performance on The Awareness of Social Inference Test: TASIT. *Brain Impairment, 16*(1), 3–18. https://doi.org/10.1017/BrImp.2015.7

McDonald, S., Flanagan, S., Martin, I., & Saunders, C. (2004). The ecological validity of TASIT: A test of social perception. *Neuropsychological Rehabilitation, 14*(3), 285–302.

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & 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. *Frontiers in Psychology, 6*, 1503.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of second-order beliefs by 5-to 10-year-old children. *Journal of Experimental Child Psychology, 39*(3), 437–471.

Poletti, M., Enrici, I., Bonuccelli, U., & Adenzato, M. (2011). Theory of Mind in Parkinson's disease. *Behavioural Brain Research, 219*(2), 342–350.

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

Repacholi, B., & Slaughter, V. (2004). *Individual differences in theory of mind: Implications for typical and atypical development.* Psychology Press.

Roca, M., Manes, F., Gleichgerrcht, E., Ibáñez, A., González de Toledo, M. E., Marenco, V., et al. (2014). Cognitive but not affective theory of mind deficits in mild relapsing-remitting multiple sclerosis. *Cognitive and Behavioral Neurology, 27*(1), 25–30.

Rosi, A., Cavallini, E., Bottiroli, S., Bianco, F., & Lecce, S. (2016). Promoting theory of mind in older adults: Does age play a role? *Aging & Mental Health, 20*(1), 22–28.

Rossetto, F., Baglio, F., Massaro, D., Alberoni, M., Nemni, R., Marchetti, A., et al. (2020). Social Cognition in Rehabilitation Context: Different Evolution of Affective and Cognitive Theory of Mind in Mild Cognitive Impairment. *Behavioural Neurology, 2020*, 5204927.

Rossetto, F., Castelli, I., Baglio, F., Massaro, D., Alberoni, M., Nemni, R., et al. (2018). Cognitive and Affective Theory of Mind in Mild Cognitive Impairment and Parkinson's Disease: Preliminary Evidence from the Italian Version of the Yoni Task. *Developmental Neuropsychology, 43*(8), 764–780.

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.

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Levkovitz, Y. (2007). The neuroanatomical basis of affective mentalizing in schizophrenia: Comparison of patients with schizophrenia and patients with localized prefrontal lesions. *Schizophrenia Research, 90*(1–3), 274–283.

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex, 46*(5), 668–677.

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., Goldsher, D., & Aharon-Peretz, J. (2005). Impaired "affective theory of mind" is associated with right ventromedial prefrontal damage. *Cognitive and Behavioral Neurology, 18*(1), 55–67.

Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012a). Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science, 23*(8), 842–847. https://doi.org/10.1177/0956797612439070

Schneider, D., Regenbogen, C., Kellermann, T., Finkelmeyer, A., Kohn, N., Derntl, B., Schneider, F., & Habel, U. (2012b). Empathic behavioral and physiological responses to dynamic stimuli in depression. *Psychiatry Research, 200*(2–3), 294–305. https://doi.org/10.1016/j.psychres.2012.03.054

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

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment, 80*(1), 99–103.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55.

Valle, A., Massaro, D., Castelli, I., & Marchetti, A. (2015). Theory of Mind Development in Adolescence and Early Adulthood: The Growing Complexity of Recursive Thinking Ability. *Europe's Journal of Psychology, 11*(1), 112–124.

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., et al. (2013). The "Reading the Mind in the Eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry, 18*(4), 326–354.

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

Yordanova, J., Kolev, V., Hohnsbein, J., & Falkenstein, M. (2004). Sensorimotor slowing with ageing is mediated by a functional dysregulation of motor-generation processes: Evidence from high-resolution event-related potentials. *Brain, 127*(Pt 2), 351–362.

Zalla, T., & Korman, J. (2018). Prior knowledge, episodic control and theory of mind in autism: Toward an integrative account of social cognition. *Frontiers in Psychology, 9*, 752. https://doi.org/10.3389/fpsyg.2018.00752