Italian validation of the mentalization scale (MentS)

Marina Cosenza[1] · Barbara Pizzini[2,3] · Mariagiulia Sacco[1] · Francesca D'Olimpio[1] · Alda Troncone[1] · Maria Ciccarelli[1] · Aleksandar Dimitrijević[4] · Giovanna Nigro[1]

Abstract

The research aimed to assess the reliability, factor structure, and validity of the Italian adaptation of the Mentalization Scale (MentS), a 28-item self-report questionnaire that measures mentalization across three dimensions. The psychometric properties of the Italian version were examined in two studies with large samples of adults and adolescents. The first study (Study 1) aimed to evaluate, through exploratory and confirmatory factor analyses, the construct validity of the Italian version of the MentS in adolescents ($N=618$) and adults ($N=720$). The second study (Study 2) was undertaken to test the convergent validity and temporal stability of the Italian version of the MentS. Specifically, the study assessed the relationship between the MentS and scores on the Reflective Functioning Questionnaire (RFQ-8), one of the most widely used instruments to assess mentalization, in a large sample of high-school students ($N=472$). Furthermore, the study evaluated the 4-week test-retest reliability of the instrument in a sample of undergraduates ($N=128$). The questionnaire exhibited strong internal consistency across both adult and adolescent samples, with Cronbach's alphas ranging from 0.71 to 0.83. Exploratory and confirmatory factor analyses consistently identified three correlated underlying factors within both age groups, demonstrating the robust factor structure of the Italian version of the MentS. Furthermore, the tool demonstrated strong convergent validity with the RFQ-8 and acceptable test-retest reliability over a 4-week period. These findings provide compelling evidence supporting the Italian version of the MentS as a reliable and valid self-report measure for comprehensively assessing different facets of mentalization.

Keywords Mentalization · Mentalization assessment · Mentalization Scale (MentS) · Reflective functioning questionnaire · Italian validation

✉ Marina Cosenza
marina.cosenza@unicampania.it

Barbara Pizzini
b.pizzini@unifortunato.eu

Mariagiulia Sacco
mariagiulia.sacco@unicampania.it

Francesca D'Olimpio
francesca.dolimpio@unicampania.it

Alda Troncone
alda.troncone@unicampania.it

Maria Ciccarelli
maria.ciccarelli@unicampania.it

Aleksandar Dimitrijević
a.dimitrijevic@ipu-berlin.de

Giovanna Nigro
giovanna.nigro@unicampania.it

1 Department of Psychology, University of Campania Luigi Vanvitelli, Viale Ellittico, 31, 81100 Caserta, Italy

2 Giustino Fortunato Telematic University, Benevento, Italy

3 Giustino Fortunato University, Viale Raffaele Delcogliano, 12, 82100 Benevento, Italy

4 International Psychoanalytic University, IPU, Stromstr. 3, 10555 Berlin, Germany

Introduction

In recent decades, mentalization (or mentalizing), also referred to as reflective functioning (RF), has emerged as a prominent empirical topic, steadily garnering heightened attention and interest. The term "mentalization" was first used in the 1960s and 1970s as a clinical psychoanalytic concept mostly related to psychosomatic states and conditions (Bion, 1962; Marty, 1991). More recently, it was defined as "the mental process by which an individual implicitly and explicitly interprets the actions of himself and others as meaningful based on intentional mental states such as personal desires, needs, feelings, beliefs, and reasons" (Bateman & Fonagy, 2004, p. XXI). This process is regarded as pivotal in both emotional and cognitive development. It is closely linked to issues such as aggressiveness, delinquency, substance abuse, and various mental disorders (for reviews, see Luyten et al., 2020, Johnson et al., 2022; see also Chevalier et al., 2023).

During the early 1990s, Peter Fonagy and his colleagues developed a tool aimed at assessing individuals' capacity to reflect on their attachment experiences. This tool, known as the Reflective Function Scale (RFS; Fonagy et al., 1998), remains the gold standard in mentalization research, renowned for its remarkable reliability, validity, and well-established structure (Tauber et al., 2013). As stressed by some authors (e.g., Hüwe et al., 2023), the RF scale is very demanding, as it requires training and certification. The Reflective Functioning Questionnaire (RFQ; Fonagy et al., 2016) was expected to address these concerns, offering promise due to its concise format of only eight items and the convenient scoring system available online.

Recent research by Müller et al. (2022) highlights significant issues with the validity of the RFQ-8. While it aims to assess individuals' ability to understand both their own and others' behavior through intentional mental states, it predominantly focuses on self-understanding rather than providing a comprehensive view of both self and others. Consequently, the RFQ-8 fails to capture the multifaceted nature of mentalization adequately. Moreover, its two subscales only address extremes of certainty and uncertainty, representing a limited portion of the broader mentalization spectrum. Further analysis reveals a unifactorial structure, lacking a bifactorial composition. In response, Horváth et al. (2023) introduced the RFQ-7, offering a streamlined questionnaire and innovative scoring system to address these limitations. This unidimensional tool includes a dimension of *hypomentalization*[1], spanning from low to high uncertainty levels.

A large body of studies demonstrated that poor mentalization is associated with several mental disorders, including borderline and antisocial personality disorders (Fonagy et al., 2016; Perroud et al., 2017), depression (Luyten et al., 2012), and eating disorders (Pedersen et al., 2015; Skårderud, 2007a, b; see also Fonagy et al., 2016). Additionally, a deficit in the capacity to "hold mind in mind" is associated, among others, with substance abuse (Allen et al., 2008; Lecointe et al., 2016; Möller et al., 2017; Suchman et al., 2018), gambling disorder (Ciccarelli et al., 2021, 2022a, b; Cosenza et al., 2019; Lindberg et al., 2011; Nigro et al., 2019; Spada & Roarty, 2015), as well as with other forms of out-of-control behaviors, such as sexual (Berry & Berry, 2014) and food addiction (Innamorati et al., 2017).

However, it remains unclear to what degree these associations specifically indicate a deficit in self-understanding rather than a broader deficiency in mentalizing. The RFQ-8, with its focus predominantly on self-awareness, save for one item, leaves ambiguity regarding whether the aforementioned connections primarily signify a lack of self-comprehension or a general impairment in mentalizing abilities (Müller et al., 2022; see also Müller et al., 2023).

Other self-report measures assessing mentalization have emerged concurrently with the Reflective Functioning Questionnaire. The first was the Mentalization Questionnaire (MZQ; Hausberg et al., 2012), a 15-item scale featuring four subscales. However, its reliability ranges between 0.57 and 0.68, and its Italian translation diverged notably from the original version, indicating a potentially unidimensional structure rather than the intended four-dimensional framework (as observed in Ponti et al., 2019). Subsequently, the Mentalization Scale (MentS; Dimitrijevic et al., 2018) gained widespread international use and consistently demonstrated robust performance with minimal observed shortcomings.

The Mentalization Scale (MentS) consists of 28 self-report items, utilizing a 5-point Likert scale from *completely agree* to *completely disagree*. Elevated scores indicate a more advanced capacity for mentalization. Typically, respondents take approximately 10 min to complete the assessment.

In the extensive community sample utilized for validation, MentS showcased robust reliability (Cronbach's Alpha = 0.84), with subscale reliabilities at 0.76 (MentS-S) and 0.77 (MentS-O and MentS-M). It demonstrated commendable whole-scale reliability and strong convergent-discriminant validity by exhibiting meaningful correlations with related constructs and fundamental personality traits. Moreover, the scale effectively differentiated between individuals with borderline personality disorder and controls,

---

[1] Hypomentalizing and hypermentalizing represent distinct impairments within RF. Hypomentalizing is characterized by an inability to engage with intricate models of one's own mind or others'. In contrast, hypermentalization involves continuous attempts to mentalize without effectively integrating cognition and emotions (Fonagy et al., 2016).

revealing significant distinctions. While the clinical sample exhibited acceptable internal consistencies across all subscale scores, MentS-M presented a deviation (for specific details, refer to Dimitrijevic et al., 2018).

Since its introduction in 2018, the MentS scale has garnered considerable attention, prompting validation studies across various linguistic contexts. These studies have encompassed translations into Chinese (Wen et al., 2022), Farsi (Ahmadian & Ghamarani, 2021), Persian, and Iranian (Ahmadian & Ghamarani, 2021; Asgarizadeh et al., 2023), as well as Japanese (Matsuba et al., 2022), Korean (Surim & Munhee, 2018), Polish (Jańczak, 2021), and Turkish (Törenli Kaya et al., 2023). Currently, efforts are underway for translations into Catalan, German, and Spanish.

Furthermore, the scale has been employed in various peer-reviewed research studies conducted across multiple languages, including French (Francoeur et al., 2020), Hindi (Bhola & Mehrotra, 2021), Hungarian (Fekete et al., 2019), Lithuanian (Gervinskaitė-Paulaitienė et al., 2023), Norwegian (Brattland et al., 2022), and Serbian (Berleković & Dimitrijević, 2020).

Several studies suggest that the MentS serves as a valid and reliable self-report tool for assessing mentalization. Its effectiveness extends to efficiently assessing sizable community samples and proving advantageous in clinical research. Across these studies, the three-factor structure of the MentS consistently emerged, with a few individual items occasionally loading on unintended factors. Test-retest reliability coefficients typically ranged from 0.68 to 0.85. Cronbach's alphas for the overall scale varied between 0.73 and 0.86, except for the Turkish version where it was 0.63. Subscale scores ranged predictably lower yet remained between 0.74 and 0.80. Notably, the correlation between scores from the Reflective Function Scale (RFS) and MentS was 0.65 ($p < 0.01$), indicating a significant relationship. Moreover, correlations between individual subscales were observed to vary between 0.41 and 0.56, remaining statistically significant across all three cases (Richter et al., 2021).

Considering the substantial evidence supporting its robustness and utility, we opted to validate the MentS in Italian. This paper presents comprehensive details of our validation study.

Overview of studies

The current research aimed to explore the psychometric properties of the Italian adaptation of the MentS through two studies. Initially, the MentS underwent translation into Italian, following the meticulous procedure outlined by Beaton et al. (2000), involving forward and backward translation, as well as pilot testing. Participants were drawn from both adult and adolescent populations.

Study 1 focused on assessing the construct validity of the Italian version of the MentS in adolescents and adults, utilizing exploratory and confirmatory factor analyses. In Study 2, the convergent validity and temporal stability of the Italian MentS were examined. Specifically, Study 2 delved into evaluating the correlation between the MentS and the Reflective Functioning Questionnaire (RFQ-8; Fonagy et al., 2016) in a substantial cohort of high-school students. Additionally, it gauged the 4-week test-retest reliability of the instrument among undergraduates. For both studies, we have reported the actual number of participants. Incomplete questionnaires (approximately 2% for both samples) were excluded from the final samples.

Consistent with the original MentS version, our expectation was to reproduce the scale's three-dimensional structure and to observe gender differences in both adult and adolescent samples. In addition, we expected significant correlation between MentS and RFQ-8 scores among the adolescent sample.

All studies were carried out in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Department of Psychology of the first author's university. Before participation, all subjects provided informed consent. For minors, informed consent was obtained from parents.

Study 1

Method

Participants

In Study 1, a total of 1338 participants from both adolescent and adult cohorts were involved. The adolescent subset encompassed 618 high school students (48.5% boys; M_{age} = 17.67; $SD = 0.53$) attending various public high schools, including lyceums and technical and trade schools in Southern Italy. These participants were randomly divided into two groups of equal size, with the first group (151 boys and 158 girls; M_{age} = 17.61; $SD = 0.70$) used for the exploratory factor analysis (EFA) and the second group (149 boys and 160 girls; M_{age} = 17.73; $SD = 0.65$) for the confirmatory factor analysis (CFA).

The adult cohort consisted of 720 volunteers (42.4% men), ranging in age from 20 to 65 years (M_{age} = 38.28; $SD = 14.69$), recruited from a community-based population. Like the adolescent group, this sample was randomly split into two equivalent groups. The first adult group (151 males and 209 females; M_{age} = 39.17; $SD = 14.36$) underwent exploratory factor analysis (EFA), while the second group (154 males and 206 females; M_{age} = 37.40 years;

SD = 14.98) participated in the confirmatory factor analysis (CFA).

Inizio modulo.

Measures

The Mentalization Scale. The MentS is structured into three distinct subscales, namely: The Self-related Mentalization subscale (MentS-S), the Other-related Mentalization subscale (MentS-O), and the Motivation to Mentalize subscale (MentS-M).

The MentS-S scale comprises eight items that center on the individual's perception of their ability to comprehend their own mental states (e.g., 18. "I find it difficult to admit to myself that I am sad, hurt, or afraid"; 22. "It is difficult for me to find adequate words to express my feelings). The MentS-O dimension consists of ten items aimed at gauging the individual's confidence in understanding the mental states of others (e.g., 10. "I can make good predictions of other people's behavior when I know their beliefs and feelings"; 20. "I can describe significant traits of people who are close to me with precision and in detail"). Finally, the MentS-M subscale encompasses ten items aimed at assessing the individual's inclination towards utilizing their capacity for mentalizing and how significant this mentalizing ability is to them (e.g., 7. "When someone annoys me, I try to understand why I react in that way"; 17. "I like reading books and newspaper articles about psychological subjects").

Statistical analyses

Data analyses were performed using IBM SPSS version 29.0. The significance threshold was set at $p < 0.05$. Initially, all variables underwent scrutiny for missing data, distribution irregularities, and outlier identification. Univariate analysis of variance (ANOVA) was employed to examine gender differences in the data.

For both the adolescent and adult samples, scores obtained from the Italian version of the MentS underwent a principal components analysis followed by Oblimin rotation with Kaiser normalization. Before conducting the analyses, three key indices, as recommended by Field (2013), were assessed to ensure the data's suitability for factor analysis. These included the Kaiser-Meyer-Olkin measure (KMO) for sampling adequacy, the determination of the correlation matrix to detect multicollinearity, and the Bartlett's test of sphericity. Bartlett's test specifically evaluates the null hypothesis that the original correlation matrix is an identity matrix (Field, 2013, p. 695). Confirmatory factor analysis was carried out utilizing the Eq. 6.2 software program

designed for structural equation modeling, as detailed by Bentler (2008).

Inizio modulo.

Results

Initially, to explore potential gender-based differences in MentS scores, univariate ANOVA was employed. As expected, results from both the adolescent and adult samples revealed noteworthy disparities: males attained significantly higher scores in the Self dimension, whereas females exhibited superior performance in the Others and Motivation dimensions, along with the overall MentS score.

Table 1 displays descriptive statistics for the entire samples as well as breakdowns by gender, along with Cronbach's alpha values and the results of the univariate ANOVA. The reliability of the MentS subscales was confirmed for both the adolescent and adult samples, as evidenced by the Cronbach's alpha values reported in Table 1.

It is important to note that we initially calculated Cronbach's alpha coefficients to facilitate comparisons with the original version of the MentS and its subsequent adaptations. However, to obtain a more refined and accurate measure of reliability, we also computed the omega coefficient (ω; McDonald, 1999) for each subscale and the total score. It is common knowledge that this coefficient goes beyond Cronbach's alpha by incorporating both item factor loadings and uniqueness, resulting in a more nuanced and precise estimation of reliability. As Table 1 shows, the values of internal consistency as measured by omega coefficients were good for the full scale and the subscales of the MentS.

Exploratory factor analysis

In both samples, the Kaiser-Meyer-Olkin (KMO) values were notably high (Adolescents = 0.822; Adults = 0.861). The determinant of the correlation matrix was consistent at 0.001 for both groups, and Bartlett's test of sphericity returned significant results (Adolescents: $\chi^2(378) = 2016.64$; $p < 0.001$; Adults: $\chi^2(378) = 2378.76$; $p < 0.001$). These outcomes signified sufficiently large correlations, supporting the suitability of the data for Principal Component Analysis (PCA).

In each case, the determination of retained factors relied on parallel analysis, conducted using the SPSS syntax developed by O'Connor (2000). Parallel analysis consistently indicated a three-component solution as the most appropriate for both samples.

In the adolescent sample, three factors collectively explained 35.95% of the variance. The first factor accounted for 18.29% of the variance, the second factor accounted for

Table 1 Study 1 – Descriptive statistics, reliabilities, and gender differences

	Samples descriptive statistics								Descriptive statistics by gender				ANOVA by gender		
									Males		Females				
	N	M	SD	Range	Skewness	Kurtosis	α	ω	M	SD	M	SD	F	p	η_p^2
Study 1															
Adolescents	618												$F_{1,616}$	p	η_p^2
MentS[a]-Self		24.56	5.44	12–40	0.071	−0.360	0.70	0.71	25.23	5.49	23.93	5.32	8.83	0.003	0.014
MentS-Others		36.59	5.26	19–50	−0.247	0.235	0.75	0.76	35.82	0.5.57	37.32	4.84	12.76	0.000	0.020
MentS-Motivation		38.13	5.50	18–50	−0.434	−0.073	0.71	0.71	36.71	5.62	39.48	5.03	41.92	0.000	0.064
MentS Total		99.29	10.96	66–136	−0.084	0.264	0.76	0.71	97.76	11.62	100.74	10.10	11.61	0.001	0.018
Adults	720												$F_{1,718}$	p	η_p^2
MentS-Self		26.53	5.25	8–40	0.067	0.126	0.71	0.72	27.22	5.05	26.03	5.34	9.18	0.010	0.013
MentS-Others		38.55	4.81	25–50	−0.201	0.530	0.78	0.79	37.39	4.76	39.40	4.67	31.97	0.000	0.043
MentS-Motivation		39.04	5.17	22–50	−0.379	0.057	0.75	0.75	37.46	4.94	40.21	5.03	53.07	0.000	0.069
MentS Total		104.13	11.63	73–138	0.134	0.163	0.84	0.83	102.07	11.59	105.63	11.75	16.81	0.001	0.023
Study 2	472												$F_{1,470}$	p	η_p^2
MentS-Self		24.59	5.77	11–40	0.242	−0.060	0.72	0.72	25.37	5.93	23.98	5.58	6.77	0.010	0.014
MentS-Others		36.26	6.30	10–50	−0.858	2.147	0.80	0.80	35.33	6.88	36.99	5.71	8.18	0.004	0.017
MentS-Motivation		37.65	6.09	18–50	−0.549	0.320	0.70	0.71	36.16	6.36	38.82	5.62	23.20	0.000	0.047
MentS Total		98.50	12.55	60–136	−0.015	0.272	0.79	0.75	96.86	13.43	99.79	11.68	6.42	0.012	0.013
RFQ-8[b] Certainty		0.90	0.72	0.00–3	0.649	−0.285	0.69	–	0.99	0.77	0.83	0.66	5.89	0.016	0.012
RFQ-8 Uncertainty		0.83	0.58	0.00–3	0.739	0.218	0.68	–	0.75	0.58	0.88	0.57	6.16	0.013	0.013

[a]Mentalization Scale; [b]Reflective Functioning Questionnaire

Table 2 Pattern matrix for the Mentalization Scale (MentS) – Adolescent sample

Item number	Component		
	MentS-O	MentS-S	MentS-M
18	0.00	**0.75**	0.03
26	−0.01	**0.73**	0.09
14	−0.03	**0.68**	−0.04
19	0.07	**0.60**	−0.21
8	0.05	**0.58**	−0.04
11	−0.03	**0.54**	0.09
22	−0.01	**0.47**	0.31
21	−0.04	**0.46**	−0.03
12	**0.73**	−0.13	−0.10
6	**0.70**	0.04	0.05
3	**0.68**	0.03	0.07
20	**0.63**	0.06	−0.09
10	**0.61**	−0.10	−0.01
5	**0.59**	0.03	0.08
23	**0.57**	−0.01	0.19
2	**0.39**	−0.06	0.17
25	**0.38**	0.35	−0.09
28	**0.32**	0.11	0.10
24	0.21	−0.09	**0.60**
27	−0.10	0.13	**0.59**
15	0.03	0.00	**0.57**
7	−0.04	−0.03	**0.56**
13	0.12	0.05	**0.53**
9	−0.05	0.18	**0.49**
1	0.17	0.00	**0.47**
4	0.28	−0.17	**0.46**
16	0.08	−0.14	**0.44**
17	−0.01	−0.06	**0.43**

Extraction method: Principal Component Analysis; Rotation method: Oblimin with Kaiser normalization. Loadings in bold indicate the highest loading across components for that item. MentS-S: MentS-Self; MentS-O: MentS-Others; MentS-M: MentS-Motivation

Table 3 Pattern matrix for the MentS – Adult sample

Item number	Component		
	MentS-O	MentS-S	MentS-M
21	0.21	**0.73**	−0.13
22	0.23	**0.72**	−0.10
8	0.06	**0.71**	−0.23
11	0.11	**0.66**	−0.10
14	−0.12	**0.48**	0.19
18	−0.12	**0.43**	0.22
19	−0.11	**0.37**	0.24
26	0.03	**0.32**	0.27
6	**0.71**	0.14	0.00
12	**0.68**	−0.08	−0.01
3	**0.63**	0.05	0.06
10	**0.61**	0.10	−0.12
23	**0.59**	0.03	0.01
5	**0.59**	−0.02	0.16
20	**0.52**	−0.07	0.12
25	**0.50**	0.29	0.00
2	**0.43**	−0.06	0.20
28	**0.30**	−0.03	0.18
9	0.02	0.20	**0.62**
24	0.24	−0.08	**0.56**
1	−0.09	−0.06	**0.54**
4	0.23	−0.17	**0.54**
7	0.14	0.06	**0.52**
27	−0.02	0.24	**0.47**
17	0.20	0.01	**0.45**
13	0.28	0.12	**0.37**
15	0.19	−0.06	**0.36**
16	0.24	−0.10	**0.31**

Extraction method: Principal Component Analysis; Rotation method: Oblimin with Kaiser normalization. Loadings in bold indicate the highest loading across components for that item. MentS-S: MentS-Self; MentS-O: MentS-Others; MentS-M: MentS-Motivation

11.55%, while the third factor explained 6.11% of the variance. Table 2 presents individual item loadings on these retained components. Notably, the first factor encompassed the ten items of the MentS Others subscale, the second factor comprised the eight items of the MentS Self scale, and the third factor consisted of the ten items of the MentS Motivation scale.

Inizio modulo.

Regarding the adult sample, the three-factor solution accounted for a cumulative variance of 36.07%. The distribution of variance across these factors was as follows: the first factor explained 21.41%, the second factor accounted for 9.27%, and the third factor elucidated 5.39%. These relationships are detailed in Table 3, which displays the subscale loadings across the three dimensions. Factor loadings distinctly revealed the composition of each factor: the first factor encapsulated the ten items of the MentS Others

subscale, the second factor encompassed the eight items of the MentS Self scale, and the third factor included the ten items of the MentS Motivation dimension.

Confirmatory factor analysis

Confirmatory Factor Analysis (CFA) utilizing maximum likelihood estimation was employed to examine the reproducibility of the proposed factor structure outlined by Dimitrijević et al. (2018) in adolescent and adult Italian samples.

In both samples, three models underwent testing. The initial model was a one-factor structure where all items were anticipated to load onto a single factor. The second model consisted of three factors with no correlation among them, while the third model allowed for intercorrelation between the factors.

Each model's goodness of fit was assessed using various measures: the likelihood ratio chi-square test statistic, adjusted for data nonnormality using Satorra and Bentler's method (1994; S-B χ^2), alongside four descriptive fit indices: standardized root-mean-square residual (SRMR), root-mean-square error of approximation (RMSEA) with its 90% confidence interval (90% CI), goodness of fit index (GFI), and comparative fit index (CFI). Considering the sensitivity of the χ^2 statistic to sample size (MacCallum, 1990; Marsh et al., 1988), interpretations of model fit were guided by a range of fit indices. Adequate model fit was identified by a non-significant S-B χ^2, GFI, and CFI values of 0.90 or higher, as well as an RMSEA less than 0.08.

Table 4 presents the model fit statistics for the three models across both groups. The model exhibiting the highest GFI and CFI estimates while displaying the lowest RMSEA and SRMR values was considered the most suitable or best-fitting model.

Discussion

The study aimed to assess the reliability of the MentS and examine its factor structure within substantial cohorts of adolescents and adults. Reliability analysis demonstrated that the MentS subscales exhibit good internal consistency. Moreover, outcomes from exploratory factor analysis notably indicated that the three-factor model appropriately captured a substantial proportion of variance, reflected in strong factor loadings.

In line with Dimitrijević et al. (2018), both exploratory and confirmatory factor analyses consistently supported a three-factor structure for the Italian version of the MentS

Table 4 Confirmatory factor analysis fit indexes for alternative models

	S-B χ^2	GFI	CFI	RMSEA [90% CI]	SRMR
Adolescents					
Unidimensional model	2.14	0.83	0.67	0.061 [0.055, 0.067]	0.078
Three uncorrelated factors	1.59	0.90	0.83	0.044 [0.037, 0.050]	0.100
Three correlated factors	1.27	0.91	0.92	0.029 [0.020, 0.037]	0.057
Adults					
Unidimensional model	2.60	0.81	0.70	0.067 [0.061, 0.072]	0.078
Three uncorrelated factors	1.97	0.89	0.82	0.052 [0.046, 0.058]	0.126
Three correlated factors	1.45	0.92	0.92	0.035 [0.028, 0.042]	0.057

S-B χ^2 = Satorra-Bentler scaled χ^2 statistic; GFI = Goodness of Fit Index; CFI = Comparative Fit Index; RMSEA = Root-Mean-Square Error of Approximation; 90% CI = 90% confidence interval for RMSEA; SRMR = Standardized Root-Mean-Square Residual

across adolescent and adult populations. Additionally, as anticipated, gender differences in MentS scores were observed in both samples. Specifically, males attained significantly higher scores in the Self dimension, while females reported higher scores on the Others and Motivation dimensions, as well as on the overall MentS score.

Study 2

Study 2 was undertaken to test the convergent validity and temporal stability of the Italian version of the MentS. A large sample of adolescents were administered the MentS and the Reflective Functioning Questionnaire (RFQ-8; Fonagy et al., 2016; Italian version for adolescents: Cosenza et al., 2019; see also, Bizzi et al., 2022). Furthermore, the test-retest reliability of the instrument was evaluated using a 4-week interval between measurements on a sample of undergraduate students.

Method

Participants

Four hundred and seventy-two adolescents (44.1% males), aged between 16 and 19 years (*Mean age* = 17.63; *SD* = 0.72), participated in this study. They were administered the Italian versions of the MentS and the RFQ-8. The RFQ-8, an eight-item self-rating questionnaire, is specifically designed to assess reflective functioning. Respondents rate items on a seven-point Likert scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The questionnaire comprises two subscales that tap into distinct mental processes: Certainty about mental states (RFQ_C) and Uncertainty about mental states (RFQ_U). Low agreement on the RFQ_C scale denotes a tendency toward excessive yet inaccurate mentalizing (*hypermentalizing*), while higher agreement signifies a more authentic mentalizing approach. Similarly, very high scores on the RFQ_U indicate a near absence of knowledge about mental states (*hypomentalizing*), whereas lower scores reflect recognition of the complexity of one's own and others' mental states, indicative of genuine mentalizing.

Zero-order correlations between the three dimensions of the MentS and the two subscales of the RFQ-8 were computed.

Additionally, a new sample of 128 undergraduates (24.2% males), aged between 20 and 29 years (*Mean age* = 21.22; *SD* = 1.56), completed the MentS twice to assess the scale's 4-week test-retest reliability.

Results

Results showed a strong positive correlation between MentS-Self and RFQ-8 Certainty scale scores ($r = 0.44$; $p < 0.001$), as well as a significant negative association between MentS-Self and RFQ-8 Uncertainty scale ($r = -0.43$; $p < 0.001$).

As for the temporal stability, the Italian version of MentS demonstrated an acceptable 4-week test-retest reliability for the three dimensions of the instrument, as well as for the full scale (MentS-Self: $r = 0.63$; $p < 0.001$; MentS-Others: $r = 0.65$; $p < 0.001$; MentS-Motivation: $r = 0.63$; $p < 0.001$; MentS full scale: $r = 0.83$; $p < 0.001$).

Discussion

Study 2 aimed to assess both the convergent validity and temporal stability of the Italian version of the MentS among adolescents and undergraduates, respectively. The results highlighted the scale's good convergent validity with the Reflective Functioning Questionnaire (RFQ-8) and demonstrated reliable (4-week) test-retest reliability.

While the RFQ-8 is designed to assess an individual's capacity to understand intentional mental states within themselves and others (Fonagy et al., 2012; Luyten et al., 2020), our study revealed a strong correlation between RFQ-8 scores and MentS-Self scores. This outcome underscores a distinct connection between reflective functioning and self-awareness, while revealing no such association with other dimensions of the MentS. These findings suggest that the RFQ-8 may particularly emphasize the comprehension of one's own mental states rather than those of others. This aligns with previous research (e.g., Dimitrijević et al., 2018; Müller et al., 2022, 2023), highlighting the significance of introspective abilities in reflective functioning. Additional support for this notion can be found in the work of Li, Carragher, and Bird (2020).

General discussion

In recent decades, mentalization (also known as mentalizing) has surged as a prominent empirical field, steadily gaining heightened attention and interest. Imbalance in the ability to perceive and interpret both the self and others' behavior in terms of intentional mental states, such as thoughts, feelings, desires, wishes, goals, and attitudes (Fonagy et al., 2012), has received significant attention over the past years (for a review, see Luyten et al., 2020).

Research exploring the significance of mentalization in psychopathology is rapidly expanding, reflecting an increasing interest in comprehending its implications. The present studies contributed to this ongoing line of research

by developing and testing an Italian version of the MentS scale, a 28-item self-report measure of mentalization.

An initial measurement study (Study 1) employing exploratory and confirmatory factor analyses on large samples of adolescents and adults revealed support for the three-correlated factors model postulated by Dimitrijević et al. (2018). Study 2 was devoted to testing the convergent validity and temporal stability of the MentS. The results obtained from a sample of adolescents demonstrated that the MentS shows a good convergent validity with the Reflective Functioning Questionnaire (RFQ-8). In addition, results from a sample of undergraduates showed that the Italian version of MentS demonstrates good test-retest reliability for the three dimensions of the instrument and the full scale.

Notably, in all studies, we observed significant differences in MentS scores due to gender. In both adult and adolescent samples, male participants scored significantly higher on the dimension MentS-Self, but significantly lower on the subscales Motivation and Other, respectively, as well as on the MentS total score. This outcome aligns seamlessly with the conclusions drawn by Dimitrijević et al. (2018), which highlighted a superior proficiency in understanding one's mental states among males. Conversely, females exhibited greater confidence in grasping the mental states of others and demonstrated their need to understand the psychic world of self and others. Our findings suggest that gender affects mentalization, albeit in a differentiated manner depending on the specific dimension under consideration. The utilization of a multidimensional measurement approach enabled the capture of crucial nuances that might have otherwise been overlooked. As emphasized by Krach et al. (2009), the longstanding hypothesis that women differ from men in their mentalizing abilities underscores the importance of employing measurement tools capable of capturing diverse facets of mentalization when evaluating gender differences.

Limitations and future research

While the current studies advanced work on instruments assessing mentalization, at least two limitations should be considered. Firstly, our studies rely on convenience samples. Secondly, the estimation of test-retest reliability was conducted on an undergraduate sample with a notably higher percentage of females than males, potentially affecting the representation of gender in the results.

Future research evaluating the extent to which the three subfactors differentially predict outcomes in substantive domains is desirable. The use of a multidimensional instrument, such as the MentS, could help in clarifying this relevant issue and test intervention strategies focused on recovering the capacity to understand others and oneself in

terms of internal mental states, always bearing in mind the various dimensions of mentalization. Furthermore, future research ought to persist in examining gender differences associated with mentalization across both normative and clinical populations, encompassing not only adolescents and adults but also older individuals, a demographic that has received limited attention in previous studies.

Data Availability The data supporting this study's findings are available from the corresponding author, Marina Cosenza, upon reasonable request. The data are not publicly available due to privacy or ethical restrictions.

Declarations

Competing interests 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.

Ethics approval Approval was obtained from the ethics committee of University of Campania "Luigi Vanvitelli". The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Consent to participate Informed consent was obtained from all individual participants included in the study.

References

Ahmadian, Z., & Ghamarani, A. (2021). Reliability and validity of Persian version of mentalization scale in university students. *Journal of Fundamentals of Mental Health*

Allen, J. G., Fonagy, P., & Bateman, A. W. (2008). Mentalizing in clinical practice. Washington, DC: American Psychiatric Press.

Asgarizadeh, A., Vahidi, E., Seyed Mousavi, P. S., Bagherzanjani, A., & Ghanbari, S. (2023). Mentalization scale (MentS): Validity and reliability of the Iranian version in a sample of nonclinical adults. *Brain and Behavior*, *13*(8), e3114. https://doi.org/10.1002/brb3.3114.

Bateman, A., & Fonagy, P. (2004). *Psychotherapy for borderline personality disorder: Mentalization based treatment*. Oxford University Press.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014.

Bentler, P. M. (2008). *EQS structural equation modeling software*. Multivariate Software.

Berleković, V., & Dimitrijević, A. (2020). Attachment and mentalization in war veterans with and without post-traumatic stress disorder. In A. Hamburger (Ed.), *Trauma, Trust, and memory* (pp. 151–159). Routledge.

Berry, M. D., & Berry, P. D. (2014). Mentalization-based therapy for sexual addiction: Foundations for a clinical model. *Sexual and Relationship Therapy, 29*(2), 245–260. https://doi.org/10.1080/14681994.2013.856516.

Bhola, P., & Mehrotra, K. (2021). Associations between countertransference reactions towards patients with borderline personality disorder and therapist experience levels and mentalization ability. *Trends in Psychiatry and Psychotherapy*, *43*(2), 116–125. https://doi.org/10.47626/2237-6089-2020-0025.

Bion, W. R. (1962). The psycho-analytic study of thinking. *International Journal of Psychoanalysis*, *43*(4–5), 306–310. https://doi.org/10.1002/j.2167-4086.2013.00030.x.

Bizzi, F., Riva, A., Borelli, J. L., Charpentier-Mora, S., Bomba, M., Cavanna, D., & Nacinovich, R. (2022). The Italian version of the reflective functioning questionnaire: Validity within a sample of adolescents and associations with psychological problems and alexithymia. *Journal of Clinical Psychology*, *78*(4), 503–516. https://doi.org/10.1002/jclp.23218.

Brattland, H., Holgersen, K. H., Vogel, P. A., Anderson, T., & Ryum, T. (2022). An apprenticeship model in the training of psychotherapy students. Study protocol for a randomized controlled trial and qualitative investigation. *PloS One*, *17*(8), e0272164. https://doi.org/10.1371/journal.pone.0272164.

Chevalier, V., Simard, V., & Achim, J. (2023). Meta-analyses of the associations of mentalization and proxy variables with anxiety and internalizing problems. *Journal of Anxiety Disorders*. https://doi.org/10.1016/j.janxdis.2023.102694. 102694.

Ciccarelli, M., Cosenza, M., Nigro, G., Griffiths, M., & D'Olimpio, F. (2022a). Gaming and gambling in adolescence: The role of personality, reflective functioning, time perspective and dissociation. *International Gambling Studies*, *22*, 161–179. https://doi.org/10.1080/14459795.2021.1985583.

Ciccarelli, M., Nigro, G., D'Olimpio, F., Griffiths, M. D., & Cosenza, M. (2021). Mentalizing failures, emotional dysregulation, and cognitive distortions among adolescent problem gamblers. *Journal of Gambling Studies*, *37*(1), 283–298. https://doi.org/10.1007/s10899-020-09967-w.

Ciccarelli, M., Nigro, G., D'Olimpio, F., Griffiths, M. D., Sacco, M., Pizzini, B., & Cosenza, M. (2022b). The associations between loneliness, anxiety, and problematic gaming behavior during the COVID-19 pandemic: The mediating role of mentalization. *Mediterranean Journal of Clinical Psychology*, *10*(1), 1–21. https://doi.org/10.13129/2282-1619/mjcp-3257.

Cosenza, M., Ciccarelli, M., & Nigro, G. (2019). The steamy mirror of adolescent gamblers: Mentalization, impulsivity, and time horizon. *Addictive Behaviors*, *89*, 156–162. https://doi.org/10.1016/j.addbeh.2018.10.002.

Dimitrijević, A., Hanak, N., Dimitrijević, A., A., & Jolić Marjanović, Z. (2018). The mentalization scale (MentS): A self-report measure for the Assessment of Mentalizing Capacity. *Journal of Personality Assessment*, *100*(3), 268–280. https://doi.org/10.1080/00223891.2017.1310730.

Fekete, K., Török, E., Kelemen, O., Makkos, Z., Csigó, K., & Kéri, S. (2019). A mentalizáció dimenziói pszichotikus zavarokban. *Neuropsychopharmacologia Hungarica*, *21*(1), 5–11.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th edition). Sage.

Fisher, R., & Fontaine, J. R. J. (2012). Methods for investigating structural equivalence. In D. Matsumo, & F. J. R. Van de Vijver (Eds.), *Cross-cultural research methods in psychology* (Vol. 215, p. 179). Cambridge University Press.

Fonagy, P., Bateman, A., & Luyten, P. (2012). Introduction and overview. In A. Bateman, & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (pp. 3–41). American Psychiatric Publishing Inc.

Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y. W., Warren, F., Howard, S., Ghinai, R., Fearon, P., & Lowyck, B. (2016). Development and validation of a self-report measure of Mentalizing: The reflective functioning questionnaire. *PloS One*, *11*(7), e0158678. https://doi.org/10.1371/journal.pone.0158678.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). Reflective Functioning Scale manual. Unpublished manuscript, London, England.

Francoeur, A., Lecomte, T., Daigneault, I., Brassard, A., Lecours, V., & Hache-Labelle, C. (2020). Social Cognition as Mediator of romantic Breakup Adjustment in Young adults who experienced Childhood Maltreatment. *Journal of Aggression Maltreatment & Trauma*, *29*(9), 1125–1142. https://doi.org/10.1080/10926771.2019.1603177.

Gervinskaitė-Paulaitienė, L., Byrne, G., & Barkauskienė, R. (2023). Mentalization-based parenting program for child Maltreatment Prevention: A pre–post study of 12-Week Lighthouse Group Program. *Children*, *10*(6), 1047. https://doi.org/10.3390/children10061047.

Hausberg, M. C., Schulz, H., Piegler, T., Happach, C. G., Klöpper, M., Brütt, A. L., Sammet, I., & Andreas, S. (2012). Is a self-rated instrument appropriate to assess mentalization in patients with mental disorders? Development and first validation of the Mentalization Questionnaire (MZQ). *Psychotherapy Research, 22*, 699–709. https://doi.org/10.1080/10503307.2012.709325.

Horváth, Z., Demetrovics, O., Paksi, B., Unoka, Z., & Demetrovics, Z. (2023). The Reflective Functioning Questionnaire-Revised-7 (RFQ-R-7): A new measurement model assessing hypomentalization. *PloS one*, *18*(2), e0282000. https://doi.org/10.1371/journal.pone.0282000.

Hüwe, L., Laser, L., & Andreas, S. (2023). Observer-based and computerized measures of the patient's mentalization in psychotherapy: A scoping review. *Psychotherapy Research*. https://doi.org/10.1080/10503307.2023.2226812.

Innamorati, M., Imperatori, C., Harnic, D., Erbuto, D., Patitucci, E., Janiri, L., Lamis, D. A., Pompili, M., Tamburello, S., & Fabbricatore, M. (2017). Emotion regulation and mentalization in people at risk for Food Addiction. *Behavioral Medicine*, *43*(1), 21–30. https://doi.org/10.1080/08964289.2015.1036831.

Jańczak, M. (2021). Polish adaptation and validation of the mentalization scale (MentS)—A self-report measure of mentalizing. *Psychiatria Polska*, *55*(6), 1257–1274. https://doi.org/10.12740/PP/125383.

Johnson, B. N., Kivity, Y., Rosenstein, L. K., LeBreton, J. M., & Levy, K. N. (2022). The association between mentalizing and psychopathology: A meta-analysis of the reading mind in the eyes task across psychiatric disorders. *Clinical Psychology: Science and Practice*, *29*(4), 423–439. https://doi.org/10.1037/cps0000105.

Krach, S., Blümel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., & Kircher, T. (2009). Are women better mindreaders? Sex differences in neural correlates of mentalizing detected with functional MRI. *BMC Neuroscience*, *10*(1), 1–11. https://doi.org/10.1186/1471-2202-10-9.

Lecointe, P., Bernoussi, A., Masson, J., & Schauder, S. (2016). La mentalisation affective de la personnalité limite addictive: Une revue de la littérature [Affective mentalizing in addictive Borderline personality: A literature review]. *L'Encephale*, *42*(5), 458–462. https://doi.org/10.1016/j.encep.2016.02.001.

Li, E. T., Carracher, E., & Bird, T. (2020). Linking childhood emotional abuse and adult depressive symptoms: The role of mentalizing incapacity. *Child Abuse & Neglect*, *99*, 104253. https://doi.org/10.1016/j.chiabu.2019.104253.

Lindberg, A., Fernie, B. A., & Spada, M. M. (2011). Metacognitions in problem gambling. *Journal of Gambling Studies*, *27*(1), 73–81. https://doi.org/10.1007/s10899-010-9193-1.

Luyten, P., Campbell, C., Allison, E., & Fonagy, P. (2020). The Mentalizing Approach to Psychopathology: State of the art and future directions. *Annual Review of Clinical Psychology*, *16*, 297–325. https://doi.org/10.1146/annurev-clinpsy-071919-015355.

Luyten, P., Fonagy, P., Lemma, A., & Target, M. (2012). Depression. In A. Bateman, & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (pp. 385–417). American Psychiatric Association.

MacCallum, R. C. (1990). The need for alternative measures of fit in Covariance structure modeling. *Multivariate Behavioral Research*, *25*(2), 157–162. https://doi.org/10.1207/s15327906mbr2502_2.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, *103*(3), 391–410. https://doi.org/10.1037/0033-2909.103.3.391.

Marty, P. (1991). *Mentalisation et psychosomatique*. Laboratoire Delagrange.

Matsuba, Y., Haraguchi, Y., Iwasaki, M., Otsuki, T., & Katsuragawa, Y. (2022). Development of the Japanese version of mentalization scale (MentS-J) and examination of its reliability and validity. *Japanese Journal of Developmental Psychology*. https://doi.org/10.11201/jjdp.33.137

McDonald, R. P. (1999). *Test theory: A unified approach*. Lawrence Erlbaum.

Möller, C., Karlgren, L., Sandell, A., Falkenström, F., & Philips, B. (2017). Mentalization-based therapy adherence and competence stimulates in-session mentalization in psychotherapy for borderline personality disorder with co-morbid substance dependence. *Psychotherapy Research*, *27*(6), 749–765. https://doi.org/10.1080/10503307.2016.1158433.

Müller, S., Wendt, L. P., Spitzer, C., Masuhr, O., Back, S. N., & Zimmermann, J. (2022). A critical evaluation of the reflective functioning questionnaire (RFQ). *Journal of Personality Assessment*, *104*(5), 613–627. https://doi.org/10.1080/00223891.2021.1981346.

Müller, S., Wendt, L. P., & Zimmermann, J. (2023). Development and validation of the Certainty about Mental States Questionnaire (CAMSQ): A self-report measure of mentalizing oneself and others. *Assessment*, *30*(3), 651–674. https://doi.org/10.1177/10731911211061280.

Nigro, G., Matarazzo, O., Ciccarelli, M., D'Olimpio, F., & Cosenza, M. (2019). To chase or not to chase: A study on the role of mentalization and alcohol consumption in chasing behavior. *Journal of Behavioral Addictions*, *8*(4), 743–753. https://doi.org/10.1556/2006.8.2019.67.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods Instruments & Computers*, *32*(3), 396–402. https://doi.org/10.3758/bf03200807.

Pedersen, S. H., Poulsen, S., & Lunn, S. (2015). Eating disorders and Mentalization: High reflective functioning in patients with Bulimia Nervosa. *Journal of the American Psychoanalytic Association*, *63*(4), 671–694. https://doi.org/10.1177/0003065115602440.

Perroud, N., Badoud, D., Weibel, S., Nicastro, R., Hasler, R., Küng, A. L., Luyten, P., Fonagy, P., Dayer, A., Aubry, J. M., Prada, P., & Debbané, M. (2017). Mentalization in adults with attention deficit hyperactivity disorder: Comparison with controls and patients with borderline personality disorder. *Psychiatry Research*, *256*, 334–341. https://doi.org/10.1016/j.psychres.2017.06.087.

Ponti, L., Stefanini, M. C., Gori, S., & Smorti, M. (2019). The assessment of mentalizing ability in adolescents: The Italian adaptation of the mentalization questionnaire (MZQ). TPM-*Testing, Psychometrics. Methodology in Applied Psychology*, *26*(1), 29–38. https://doi.org/10.4473/TPM26.1.2.

Richter, F., Steinmair, D., & Löffler-Stastka, H. (2021). Construct validity of the mentalization scale (MentS) within a mixed psychiatric sample. *Frontiers in Psychology*, *12*, 608214. https://doi.org/10.3389/fpsyg.2021.608214.

Skårderud, F. (2007a). Eating one's words, part I: 'Concretised metaphors' and reflective function in anorexia nervosa–an interview study. *European Eating Disorders Review.,15*(3), 163–174. https://doi.org/10.1002/erv.777.

Skårderud, F. (2007b). Eating one's words, part II: The embodied mind and reflective function in anorexia nervosa–theory. *European Eating Disorders, 15*(4), 243–252. https://doi.org/10.1002/erv.778.

Spada, M. M., & Roarty, A. (2015). The relative contribution of metacognitions and attentional control to the severity of gambling in problem gamblers. *Addictive Behaviors Reports*, *1*, 7–11. https://doi.org/10.1016/j.abrep.2015.02.001.

Suchman, N. E., DeCoste, C., Borelli, J. L., & McMahon, T. J. (2018). Does improvement in maternal attachment representations predict greater maternal sensitivity, child attachment security and lower rates of relapse to substance use? A second test of mothering from the Inside out treatment mechanisms. *Journal of Substance Abuse Treatment*, *85*, 21–30. https://doi.org/10.1016/j.jsat.2017.11.006.

Surim, L., & Munhee, L. (2018). Validation of the Korean Version of the mentalization scale. *Counseling Studies*, *19*(5), 117–135. https://doi.org/10.15703/kjc.19.5.201810.117.

Taubner, S., Hörz, S., Fischer-Kern, M., Doering, S., Buchheim, A., & Zimmermann, J. (2013). Internal structure of the reflective functioning scale. *Psychological Assessment*, *25*(1), 127–135. https://doi.org/10.1037/a0029138.

Törenli Kaya, Z., Alpay, E. H., Türkkal Yenigüç, Ş., & Özçürümez Bilgili, G. (2023). Validity and reliability of the Turkish version of the mentalization scale (MentS). Zihinselleştirme Ölçeği'nin Türkçe Çevirisinin Geçerlik ve Güvenirlik Çalışması. *Turk Psikiyatri Dergisi = Turkish Journal of Psychiatry*, *34*(2), 118–124. https://doi.org/10.5080/u25692.

Wen, Y., Fang, W., Wang, Y., Du, J., Dong, Y., Zu, X., & Wang, K. (2022). Reliability and validity of the Chinese version of the mentalization scale in the general population and patients with schizophrenia: A multicenter study in China. *Current Psychology*. https://doi.org/10.1007/s12144-022-04093-9.