Bilişsel Davranışçı Psikoterapi ve Araştırmalar Dergisi
Journal of Cognitive-Behavioral Psychotherapy and Research
DOI: 10.14744/JCBPR.2025.15551
J Cogn Behav Psychother Res 2025;14(4):253–263

Adaptation and Validation of the Interactive Mentalization Questionnaire in Türkiye

Buket Ünver

Department of Psychology, Işık University, Faculty of Economics, Administrative and Social Sciences, İstanbul, Türkiye

ABSTRACT

The Interactive Mentalization Questionnaire (IMQ) is the only brief, self-report instrument that assesses three complementary facets of mentalizing in social interaction: perspective-taking (self–other), metacognition (self–self), and meta-mentalization (other–self). As there is no multidimensional mentalization scale available for Turkish speakers yet, we used a multistep, forward- and back-translation procedure to translate the IMQ and examined its psychometric properties in a community sample of Turkish adults (n=953; 43% women, mean age=31.8 years). After removing four poorly performing items, exploratory and confirmatory factor analyses replicated the original three-factor structure, yielding a 16-item version with an excellent global fit, strong internal consistency (ω_t=0.87), and informative item parameters from graded-response modeling. Construct validity was supported by convergent correlations with the Turkish Mentalization Scale (MentS), prefrontal-function indices, and divergent correlations with borderline and psychopathic traits. Strict measurement invariance was established across gender, psychiatric diagnosis status, and continuous age. Reliable change indices and minimal important differences were established to facilitate clinical monitoring. Thus, the Turkish IMQ is a concise, psychometrically solid tool for research, cross-cultural comparison, and routine outcome assessment in Turkish mental health settings.

Keywords: Cultural adaptation, mentalization, meta-mentalization, perspective-taking, scale validation.

ÖZ

Etkileşimsel Zihinselleştirme Ölçeğinin Türkçe Uyarlama ve Geçerlilik Çalışması

Etkileşimsel Zihinselleştirme Ölçeği (EZÖ), sosyal etkileşimde zihinselleştirmenin üç tamamlayıcı yönünü değerlendiren tek kısa, öz bildirim aracıdır: Perspektif alma (Ben–Öteki), üstbiliş (Ben–Ben) ve meta-zihinselleştirme (Öteki–Ben). Türkçede çok boyutlu bir zihinselleştirme ölçeği bulunmadığından, EZÖ çok aşamalı, çeviri-geri çeviri yöntemiyle Türkçeye uyarlandı ve Türk yetişkinlerden oluşan bir toplum örnekleminde (n=953; %43 kadın, yaş ortalaması=31,8) psikometrik özellikleri incelendi. Açımlayıcı ve doğrulayıcı faktör analizleri, düşük performans gösteren dört madde çıkarıldıktan sonra özgün üç faktörlü yapıyı doğruladı; böylece mükemmel model uyumu ve güçlü iç tutarlılık (ω_t=0,87) gösteren 16 maddelik bir versiyon elde edildi. Zihinselleştirme ölçeği ve kişiler arası nörobiyoloji temelli prefrontal işlev indeksleriyle yakınsak korelasyonlar; borderline ve psikopatik eğilimlerle ayrışan korelasyonlar yapı geçerliliğini destekledi. Cinsiyet, psikiyatrik tanı durumu ve yaş değişkeni açısından ölçüm değişmezliği sağlandı. Klinik izlemeyi kolaylaştırmak için güvenilir değişim indeksleri ve minimal önemli fark değerleri belirlendi. Bu bulgular doğrultusunda, Türkçe EZÖ; araştırmalarda, kültürler arası karşılaştırmalarda ve klinik değerlendirme süreçlerinde kullanılabilecek kısa, geçerli ve güvenilir bir araç sunmaktadır.

Anahtar Kelimeler: Kültürel uyarlama, zihinselleştirme, meta-zihinselleştirme, perspektif alma, ölçek geçerliliği.

Cite this article as:
Ünver B. Adaptation and Validation of the Interactive Mentalization Questionnaire in Türkiye. J Cogn Behav Psychother Res 2025;14(4):253–263.

Address for correspondence:
Buket Ünver.
Işık Üniversitesi, İktisadi, İdari ve Sosyal Bilimler Fakültesi, Psikoloji Bölümü, İstanbul, Türkiye
Phone: +90 444 0 799
E-mail:
buket.unver@isikun.edu.tr

Submitted: 09.07.2025
Revised: 09.09.2025
Accepted: 01.10.2025
Available Online: 28.11.2025

JCBPR, Available online at http://www.jcbpr.org/

INTRODUCTION

Mentalization is the process by which we attribute mental states to our own and others' behavior (Fonagy & Bateman, 2019). It forms part of a broader set of socio-cognitive skills aimed at perceiving, interpreting, and processing social information from the environment. Metacognition (i.e., second-order capacity to monitor, evaluate, and regulate one's own cognitive processes) encompasses knowledge of personal beliefs, awareness of current mental states, and confidence judgments (Drigas & Mitsea, 2020; Fiedler et al, 2019; Rouault et al, 2018). By generating and flexibly updating mental representations of self–other relationships, mentalization guides socially adaptive behavior and supports the development of a coherent sense of self (Fonagy et al, 2002; Luyten & Fonagy, 2018).

Mentalization underpins relational reciprocity, emotion regulation, and empathy because it involves meaning-making and perspective-taking (Arioli et al, 2021; Majdandžić et al, 2016). Meta-mentalization (i.e., the estimation of how much insight another agent has into their own thoughts and intentions) combines perspective-taking with metacognitive appraisal and is crucial for strategic interaction (Bhatt et al, 2010; Silston et al, 2018; Wu et al, 2020). Longitudinal studies have reported positive associations between mentalization and language abilities, as well as between executive and emotional control (Osterhaus & Bosacki, 2022). Understanding others' current mental state allows individuals to anticipate future reactions based on others' beliefs (Jara-Ettinger et al, 2015; Zhang et al, 2025) and plan future actions using past observations (Ho et al, 2022), thereby extending the impact of mentalization beyond the "here and now." Deficits in mentalization are implicated in a wide range of psychiatric disorders, including social anxiety disorder, post-traumatic stress disorder, obsessive–compulsive disorder, borderline personality disorder, psychosis, autism spectrum disorder, and suicidal behavior (Bora, 2021; Johnson et al, 2022; Sloover et al, 2022).

Although mentalization plays a key role in mental health and is intertwined with various cognitive and emotional functions, it is challenging to develop a valid and reliable tool to gauge one's own and others' mentalizing capacities. Fonagy et al. (1998) introduced the RFS, which assesses individuals' capacity to reflect on their attachment experiences by coding responses to the Adult Attachment Interview (George et al, 1996). Renowned for its high reliability, validity, and factorial integrity, the RFS has become the gold standard for measuring mentalization (Taubner et al, 2013). However, its application requires extensive training and certification for interviewers and a considerable investment of time for conducting interviews, transcribing responses, and coding. These characteristics make RFS assessment complex and time-intensive, especially in large cohorts. To address these limitations, several self-report instruments, including the Mentalization Questionnaire (MZQ; Hausberg et al, 2012), the Mentalization Scale (MentS; Dimitrijević et al, 2018; Stefana et al, 2024), and the Reflective Functioning Questionnaire (RFQ; Fonagy & Bateman, 2019), have been developed. Although these scales measure both mentalization and metacognitive aspects, they do not assess meta-mentalization, a construct necessary for a comprehensive analysis of mentalization. Therefore, to fill this gap, Wu et al. (2022) recently developed and validated the Interactive Mentalizing Questionnaire (IMQ), which evaluates three interrelated components of mentalization in social interaction.

The main purpose of this study was to validate the Turkish version of the IMQ. Analyses include descriptive statistics and item properties based on classical test theory, factor structure using exploratory and confirmatory factor analysis, measurement invariance, item response theory, internal consistency, and correlations with external variables.

METHOD

Participants and the Procedure

A total of 953 Turkish adults (43% female; M=31.75 years, SD=12.16) were recruited through snowball sampling in Türkiye. Most participants reported university education (75%) and middle socioeconomic status (60%); 16% disclosed a past psychiatric or neurological diagnosis, and 10% reported current psychotropic medication. Ethical approval was obtained from the Işık University Ethics Review Board (approval no: 2024/03, date: April 16, 2024). This study was conducted in accordance with the Declaration of Helsinki to ensure the ethical standards and the rights of the participants.

Translation Procedure

Three bilingual psychologists produced forward translations following the Mapi Research Trust guidelines; a reconciliation panel created a consensus version. Two linguists ensured clarity. Ten adults piloted the draft, prompting minor lexical changes. Independent back-translation was approved by the original IMQ authors.

Measures

IMQ

The IMQ (Wu et al., 2022) is a 20-item self-report measure of mentalization. Each item is rated on a 4-point Likert scale (1=very true; 4=very false) that yields three distinct subscale scores: (i) self–other mentalization (IMQ_SO) assesses

how individuals infer others' mental states from their own perspective (e.g., "When I watch a movie, I can always guess what the character will do next"), (ii) self–self metacognition (IMQ_SS) captures one's awareness and evaluation of their own mental states (e.g., "When I fail, I know exactly why I failed"), and (iii) other–self meta-mentalization (IMQ_OS) reflects how well individuals believe others can infer their own thoughts and feelings (e.g., "Sometimes, I think people have direct insight into what I am thinking"). Subscale scores were calculated by summing the relevant items, with higher totals indicating stronger tendencies in the targeted dimension. In the original validation sample, the internal consistency was good (α=0.76 for IMQ_SO, 0.83 for IMQ_SS, and 0.81 for IMQ_OS; Wu et al, 2022).

MentS

The Turkish adaptation of the MentS (Törenli Kaya et al, 2023) comprises 25 items that provide an overall mentalization score and three subscale scores: self-related (MentS-S), other-related (MentS-O), and motivation (MentS-M). In this sample, McDonald's ω_t was 0.91 for the full scale and 0.83 for MentS-S, 0.88 for MentS-O, and 0.77 for MentS-M.

BPQ

The Turkish adaptation of the Borderline Personality Questionnaire (BPQ; Ceylan et al, 2017) is an 80-item self-report measure of borderline personality traits. It assesses nine borderline personality disorder DSM criteria (American Psychiatric Association, 2000): impulsivity, affective instability, abandonment, relationships, self-image, suicide/self-mutilation, emptiness, intense anger, and psychosis. Higher scores reflect greater borderline features. In the current sample, ω_t was 0.98 for the total scale and ranged between 0.79 and 0.91 across the nine subscales.

LSRP

The Turkish adaptation of the Levenson Self-Report Psychopathy Scale (LSRP; Engeler & Yargıç, 2004) is a 26-item self-report questionnaire developed to assess psychopathic traits, mapped onto the two-factor structure of the Psychopathy Checklist–Revised (PCL–R). The primary psychopathy subscale (16 items) paralleled the PCL–R Factor 1 (affective–interpersonal features), and the secondary-psychopathy subscale (10 items) paralleled the PCL–R Factor 2 (antisocial lifestyle). The subscale scores are summed separately and combined for a total score. In our sample, the internal consistency was high for the total scale (ω_t=0.88) and primary subscale (ω_t=0.90) and acceptable for the secondary subscale (ω_t=0.67).

IPNB-PFCFS

The Interpersonal Neurobiology–Based Prefrontal Cortex Functions Scale (IPNB-PFCFS; Hisli-Şahin & Varlık-Özsoy, 2017) is a 40-item self-report instrument designed to assess prefrontal cortex–mediated functions within an interpersonal neurobiology framework. Five subscales were scored by summing their respective items: body regulation; life and fear modulation; empathy and response flexibility; insight; and morality. A total score is obtained by summing all the items. In our sample, ω_t was 0.94 for the total score and ranged from 0.79 to 0.84 for all subscales, except body regulation (ω_t=0.60).

Statistical Analyses

Analyses were conducted using R 4.4.2. The sample was randomly divided into exploratory (n=500) and confirmatory (n=453) subsamples. EFA used principal-axis factoring with PROMAX rotation on polychoric correlations; items with primary loadings of <0.32 or cross-loadings of >0.30 were removed. CFA compared the one-factor, correlated three-factor, and bi-factor models (WLSMV estimation). Reliability was estimated using McDonald's ω_t and Cronbach's α. Graded-response IRT models were used to provide item parameters. Configural, metric, scalar, and strict invariance were tested across gender (ordinal indicators) and psychiatric status (continuous indicators), whereas age invariance was assessed using the MIMIC model. The convergent and divergent validity was evaluated using Bonferroni-corrected Pearson correlations.

RESULTS

Item Analysis

Item means ranged from 2.24 to 3.48; skewness=-1.27 to 0.23; and kurtosis=-1.15 to 1.36. Most items demonstrated moderate correlations (r≥0.30), but four items (7, 12, 13, and 18) showed poor discrimination (corrected r<0.10). Appendix 1 details the descriptive statistics for the IMQ items. Appendix 2 presents the demographic and clinical characteristics of the confirmatory, exploratory, and full samples.

Factor Structure

Parallel analysis suggested four factors when using PCA, seven factors when using squared multiple correlations, and five factors when using exploratory factor analysis. Given the limited number of items, five- or six-factor solutions were unlikely. Therefore, EFA was conducted on the exploratory subsample, extracting four correlated factors and evaluating these for adequate indicators (≥three items with loadings of ≥0.32) and conceptual coherence. Because only two items loaded above 0.32 on the fourth factor, we ran a subsequent EFA to extract three factors. After removing items 7 and 12 due

Table 1. Item response theory parameters and information statistics for the IMQ subscales (n=953)

Scale	Item	Loading	α	β_1	β_2	β_3	Total information	Peak information	Peak θ
Self–self									
	2	0.698	1.64	-3.39	-1.81	0.14	41.06	0.74	-1.9
	11	0.729	1.75	-3.12	-1.65	0.34	44.53	0.84	-1.8
	14	0.536	1.12	-4.13	-2.54	-0.34	24.28	0.37	-3.1
	15	0.627	1.46	-3.46	-1.66	0.52	36.44	0.58	-1.8
	16	0.687	1.75	-3.77	-1.78	0.30	46.87	0.80	-1.8
	17	0.272	0.49	-5.41	-1.90	0.97	7.15	0.07	-1.6
	19	0.553	1.18	-3.92	-1.59	0.98	28.95	0.38	-1.7
	20	0.756	2.10	-2.79	-1.59	0.05	52.94	1.21	-1.7
Self–other									
	3	0.867	3.26	-2.31	-1.23	0.41	91.10	2.74	-1.3
	4	0.919	4.84	-2.30	-1.13	0.36	143.04	5.87	-2.3
	5	0.460	0.88	-4.88	-1.55	1.69	20.06	0.21	-1.5
	8	0.412	0.75	-5.41	-2.29	0.92	15.13	0.16	-2.4
	10	0.665	1.64	-2.92	-1.32	0.68	41.49	0.74	-1.4
Other–self									
	1	0.680	1.14	-3.46	-1.67	-0.21	23.53	0.39	-1.3
	6	0.356	2.56	-1.75	-0.36	1.24	70.03	1.71	-0.4
	9	0.635	1.43	-2.71	-0.54	2.14	37.55	0.54	-0.6

IMQ: Interactive Mentalization Questionnaire; α: Discrimination parameter; β_1–β_3: Difficulty thresholds; Loadings: Standardized factor loadings from the confirmatory factor analysis; Peak θ: Ability level (θ) at which information is maximal.

Table 2. Descriptive statistics, measurement error, and reliable change indices for total and subscale IMQ scores (n=953)

Scale	Mean (SD)	Trimmed mean	MAD	Min–Max	Skewness	Kurtosis	SE	SE_m	SE_d	90% CC	95% CC	MID	MCRC
Total score	52.74 (5.75)	52.98	5.93	24–68	-0.57	1.21	0.19	2.08	2.93	3.41	4.06	2.88	5.75
Self–self	26.04 (3.57)	26.25	2.96	10–32	-0.62	0.56	0.12	1.51	2.14	2.49	2.97	1.79	4.20
Self–other	15.62 (2.54)	15.73	2.96	5–20	-0.49	0.53	0.08	1.14	1.61	1.87	2.23	1.27	3.15
Other–self	8.72 (1.77)	8.82	1.48	3–12	-0.56	0.19	0.06	0.95	1.35	1.57	1.87	0.89	2.64

IMQ: Interactive Mentalization Questionnaire; 90% CC: Critical change at the 90% confidence level; 95% CC: Critical change at the 95% confidence level; MAD: Median absolute deviation; MCRC: Minimum change for a reliable change; MID: Minimally important difference; SE: Standard error; SEm: Standard error of measurement; SEd: Standard error of difference; SD: Standard deviation; Trimmed Mean: Mean after trimming 10% of scores at each tail.

to high cross-loadings and items 13 and 18 due to primary loadings below 0.32, we conducted a final EFA to extract three correlated factors. Items demonstrated adequate to strong loadings on their respective factors (all loadings ≥0.33), and the item–factor distribution perfectly matched that of the original IMQ version (Wu et al, 2022). Appendix 3 and 4 present the item-level EFA results. The final 16-item Turkish version of the IMQ is presented in Appendix 5.

The one-, three-, and bi-factor models were tested using the CFA. The one-factor model of the IMQ showed poor results: χ^2/df=8.03, CFI=0.79, TLI=0.75, RMSEA=0.13, and SRMR=0.10. The three-factor model showed a good fit for the data: χ^2/df=2.26, CFI=0.96, TLI=0.96, RMSEA=0.05, and SRMR=0.06. Finally, the bi-factor model showed the best fit indices: χ^2/df=2.31, CFI=0.97, TLI=0.95, RMSEA=0.05, and SRMR=0.05.

Item Response Theory

Table 1 presents the item parameters. For self–self mentalization, the discrimination parameters ranged from 0.49 (item 17) to 2.10 (item 20), with six of eight items exceeding α=1.15, indicating a strong ability to distinguish among respondents across trait levels. Item information values varied between 7.15 and 52.94 and peaked at θ≈–3.1 to -1.6, suggesting that Factor 1 is the most precise for below-average respondents, yet it maintains acceptable precision across the continuum. For the self–other mentalization, the discrimination parameters were more variable (α=0.75–4.84), with two items (item 3, item 4) exhibiting very high sensitivity (α>3.0). The total information per item ranged from 15.13 to 143.04, peaking at θ≈–2.3. This pattern indicates that Factor 2 measures most precisely at low-to-moderate trait levels but is less sensitive at extreme levels. In contrast, Factor 3 items showed moderate discrimination (α=1.14–2.56) and lower information (23.53–70.03) with peaks between θ≈–1.3 and -0.4. These characteristics imply that Factor 3 provides limited precision for distinguishing individuals reliably, particularly at higher trait levels. Figure 1 shows the item characteristic curves illustrating these parameter differences.

Reliability Analysis

Internal consistency for the IMQ was strong: McDonald's ω_t =0.87 and Cronbach's α=0.76 for the total scale; ω_t=0.82 (α=0.74), ω_t=0.80 (α=0.73), and ω_t=0.71 (α=0.62) for Factors 1–3, respectively.

Measurement Error and Reliable Change Indices

Table 2 presents the descriptive statistics, measurement error, and reliable change indices for the total and subscale scores of the IMQ.

Correlations Between the Total and Subscale Scores

IMQ total score was very strongly associated with self–self mentalization subscore (r=0.85, p<0.001), strongly associated with self–other mentalization subscore (r=0.77, p<0.001), and weakly associated with other–self mentalization subscore (r=0.31, p<0.001). Self–self and self–other mentalization subscores were moderately correlated (r=0.50, p<0.001), whereas correlations involving other–self mentalization subscore were negligible and nonsignificant (self–self and other–self: r=0.06, p=0.092; self–other and other–self: r=–0.03, p=0.333).

Associations of Total and Subscale Scores with Sociodemographic and Clinical Variables

Age was positively and weakly associated with the self–self mentalization subscale (r=0.22, p<0.001). All other correlations between IMQ total or subscale scores and sociodemographic

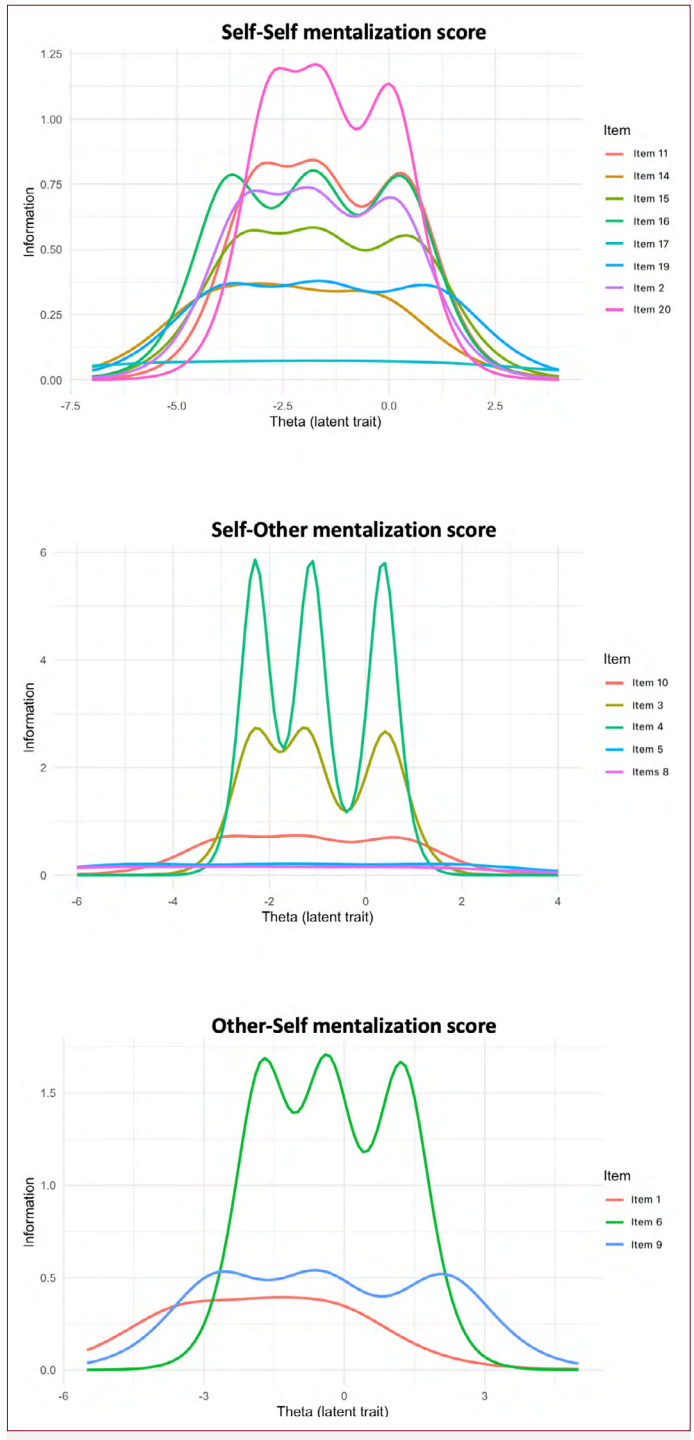

Figure 1. Item information functions.

The horizontal axis (θ) represents the underlying latent trait continuum (low to high), and the vertical axis indicates the amount of information (i.e., precision) each item provides at each trait level. The peaks of the curves show where an item is most informative (smallest standard error), and the width of each curve reflects the range of θ over which the item contributes useful information. Clear, orderly peaks and intersections indicate good discrimination and well-functioning response categories.

Table 3. Measurement invariance of the IMQ across age, gender, and psychiatric diagnosis (n=953)

Invariance	χ^2	of	CFI	ΔCFI	TLI	ΔTLI	RMSEA	ΔRMSEA	SRMR	ΔSRMR
Gender										
Configural	511.65	202	0.959	—	0.951	—	0.057	—	0.059	—
Metric	496.83	215	0.963	0.004	0.958	0.007	0.053	-0.004	0.061	0.002
Scalar	536.52	244	0.961	-0.002	0.962	0.011	0.050	-0.007	0.060	0.001
Strict	536.52	244	0.961	-0.002	0.962	0.011	0.050	-0.007	0.060	0.001
Psychiatric diagnosis										
Configural	464.01	202	0.914	—	0.898	—	0.052	—	0.049	—
Metric	478.17	215	0.913	-0.001	0.903	0.005	0.051	-0.001	0.051	0.002
Scalar	496.92	228	0.912	-0.002	0.907	0.009	0.050	-0.002	0.052	0.003
Strict	532.06	244	0.905	-0.009	0.907	0.009	0.050	-0.002	0.054	0.005
Age	441.57	114	0.956	—	0.954	—	0.055	—	0.051	—
MIMIC model										

Δ values reflect the change from the configural model within each grouping variable. Δ: Delta; MIMIC: Multiple indicators multiple causes; CFI: Comparative Fit Index; ΔCFI: Delta Comparative Fit Index; TLI: Tucker–Lewis Index; ΔTLI: Delta Tucker–Lewis Index; RMSEA: Root Mean Square Error of Approximation; ΔRMSEA: Delta root mean square error of approximation; SRMR: Standardized root mean square residual; ΔSRMR: Delta standardized root mean square residual.

(gender, educational level, socioeconomic status) or clinical (psychiatric diagnosis, psychiatric medication) variables were non–significant (rs ranged from -0.07 to 0.09, ps≥0.06).

Measurement Invariance

Table 3 shows a strict invariance across gender (ΔCFI≤0.002; ΔRMSEA≤0.003) and psychiatric status (ΔCFI≤0.004). The MIMIC model showed no differential item functioning across age (χ^2[114]=441.57, CFI=0.956, RMSEA=0.055).

Construct and Convergent Validity

As detailed in Table 4, the IMQ total score and self–self (metacognition) subscale evidenced moderate positive correlations with several construct and convergent validity measures. Specifically, the IMQ total score correlated moderately with the MentS total score (r=0.43, p<0.001), MentS-O (r=0.50, p<0.001), the IPNB-PFCFS total score (r=0.44, p<0.001) and its subscores Life and Fear Modulation subscale (r=0.44, p<0.001) and Empathy and Response Flexibility (r=0.40, p<0.001). The IMQ self–self subscale similarly showed moderate associations with MentS-O (r=0.44, p<0.001), IPNB-PFCFS total score (r=0.49, p<0.001), IPNB-PFCFS Life and Fear Modulation (r=0.46, p<0.001), and IPNB-PFCFS Empathy and Response Flexibility (r=0.45, p<0.001). All other correlations were weak (rs<0.40).

DISCUSSION

This study examines the psychometric properties of the IMQ's first Turkish translation and adaptation in a large community sample.

Summary of the Key Findings

The Turkish IMQ provides a dependable overall mentalization score plus three practical domains (self–self, self–other, other–self) consistent with the original. After removing weaker items, both the total and subscale scores performed well: total, self–self, and self–other were reliable and most informative for patients with low-to-average mentalization, whereas other–self was acceptable but less precise, especially at higher ability. Scores were similar across gender and psychiatric status, with no age-related item bias. As expected, the total and self–self/self–other scores showed moderate links to external measures of empathy, emotion regulation, and prefrontal functioning. Clinically, the total, self–self, and self–other scores can be relied on for screening and monitoring; the other–self scores are interpreted with caution, and the reliable change thresholds in Table 2 are used to judge meaningful improvement or worsening.

Structure and Hierarchy

The factor analyses supported the theorized three-factor solution. The four items were eliminated because of low item–total correlations and substantial cross-loadings. Item 18 ("Do you believe in telepathy?") was particularly problematic: its corrected item–total correlation was zero, and qualitative feedback indicated that the Turkish wording suggested a concrete perceptual ability rather than the intended meta-mentalization construct, thereby obscuring its meaning. To address such ambiguity in future adaptations, problematic items—particularly those referencing concepts such as "telepathy" or "mind reading"—should be rephrased to avoid

Table 4. Correlations between IMQ total and subscale scores and validity measures (n=953)

Measure	Total score	Self-self	Self-other	Other-self
BPQ				
Total score	-0.19†	-0.29†	0.03	-0.19†
Impulsivity	-0.01	-0.09	0.10	-0.15†
Affective instability	-0.17†	-0.28†	0.01	-0.13*
Abandonment	-0.17†	-0.22†	-0.01	-0.17†
Relationships	-0.13*	-0.20†	0.00	-0.09
Self-image	-0.27†	-0.34†	-0.10	-0.14*
Suicide/self-mutilation	-0.10	-0.19†	0.02	-0.07
Emptiness	-0.20†	-0.27†	-0.03	-0.14*
Intense anger	-0.11	-0.20†	0.03	-0.10
Psychotic states	0.06	0.01	0.24†	-0.26†
IPNB-PFCFS				
Total score	0.44†	0.49†	0.29†	0.10
Body regulation	0.28†	0.32†	0.17†	0.05
Life and fear modulation	0.44†	0.46†	0.34†	0.05
Morality	0.20†	0.31†	0.09	0.03
Empathy and response flexibility	0.40†	0.45†	0.25†	0.07
Insight	0.31†	0.30†	0.22†	0.10
LSRP				
Total score	-0.10	-0.19†	0.07	-0.23†
Primary psychopathy	-0.04	-0.13*	0.09	-0.21†
Secondary psychopathy	-0.16†	-0.23†	-0.01	-0.16†
MentS				
Total score	0.43†	0.38†	0.29†	0.26†
Motivation to mentalize	0.20†	0.14*	0.17†	0.14*
Other-related mentalization	0.50†	0.44†	0.39†	0.17†
Self-related mentalization	0.27†	0.27†	0.11	0.25†

BPQ: Borderline Personality Questionnaire; IMQ: Interactive Mentalizing Questionnaire; IPNB-PFCFS: Interpersonal Neurobiology–Based Prefrontal Cortex Functions Scale; LSRP: Levenson Self-Report Psychopathy Scale; MentS: Mentalization Scale; †: P<0.001; *: P<0.05.

misinterpretation as magical thinking. For example, phrasing that emphasizes subjective inference (e.g., "Have you ever felt that someone intuitively understood your thoughts or emotions?") may better convey the meta-mentalization construct without triggering culturally loaded associations. Incorporating such changes may enhance the face validity of the scale and reduce the risk of construct-irrelevant variance due to cultural misunderstandings.

The deletion of these items sharpened the factor structure without narrowing the content coverage. The three Turkish factors are correlated yet distinct in the original IMQ, and the final item–factor pattern is identical to that reported by Wu et al. (2022). Metacognition (self–self) showed a moderate positive association with perspective-taking (self–other), replicating earlier findings and dovetailing with simulation theory, which proposes that people rely on the same metacognitive apparatus to represent both their own and others' mental states (Carruthers, 1996; Harris, 1992). Mirror-neuron research demonstrates overlapping activation during action execution and observation (Gallese & Goldman, 1998; Gallese et al, 2007).

In contrast, correlations involving meta-mentalization (other–self) were negligible, diverging from the modest associations observed in the original validation but consistent with the initial expectations of the scale developers (Wu et al, 2022). One plausible explanation is that the ambiguous wording of the discarded items (all drawn from the self–other and other–self subscales) may have inflated the perceived ability. Notably, the original IMQ was not subjected to cognitive interviews with representatives of the target population, which is a recommended step in scale development for detecting such ambiguities (Stefana et al, 2025).

Pruning left the other–self subscale with only three items (1, 6, and 9) and highlighted two culture-specific concerns. First, the remaining items ask whether strangers can "read" or "sense" hidden feelings, language that can evoke magical thinking or paranoid interpretations rather than the belief that others sometimes grasp one's inner states. Second, interpersonal disclosure norms in Türkiye are comparatively cautious; limited epistemic trust may lead respondents to reject the premise that outsiders truly understand them, thereby lowering the mean scores and inflating the error variance. Consequently, low other–self scores likely reflect genuine doubt about being understood and measurement mismatch. This subscale should be interpreted with caution and treated as exploratory. Although the inclusion of this dimension aligns with the theoretical structure of the original IMQ, its weak psychometric performance in the Turkish adaptation suggests that it may function as a distinct or culturally constrained construct. As such, the current findings regarding the other–self factor should be viewed as preliminary and hypothesis-generating. Retaining the subscale provides a basis for theoretical continuity; however, its limited reliability and inconsistent external associations indicate the need for further refinement. Future research should include qualitative methods and cognitive interviewing to generate culturally resonant items that better capture the intended construct in Turkish populations. Subsequent pilot testing using IRT and differential item functioning analyses can refine and validate the revised items.

Measurement Invariance

To date, no study has examined the measurement invariance for the IMQ, including the original validation by Wu et al. (2022); thus, the present analyses offer the first evidence on this point. The Turkish IMQ satisfied strict invariance across gender and psychiatric diagnosis status. Changes in global fit indices indicated that factor loadings, item thresholds, and residual variances were equivalent for men and women and for participants with and without a self-reported psychiatric condition. These findings exceed the "partial measurement invariance" often reported in clinical and cross-cultural research (Leitgöb et al, 2023). A MIMIC model was used to assess age-related measurement invariance. The model demonstrated that item functioning is unbiased across the adult lifespan represented in our sample (18–88 years). Taken together, these results provide strong evidence that the Turkish IMQ yields scores that are directly comparable across gender, mental health status, and age, granting researchers and clinicians confidence that the observed group differences/similarities in mentalization are substantive rather than artifactual.

Associations with External Constructs

The correlational pattern was broadly consistent with the theoretical expectations. The overall IMQ score, particularly the self–self mentalization subscale, showed moderate positive links with both the MentS and the IPNB-PFCFS. These associations reinforce the idea that monitoring one's own mental states is embedded in a wider network of empathic and regulatory skills. Metacognition (self–self) was most strongly related to the IPNB-PFCFS total score and its Life and Fear Modulation subscale and Empathy and Response Flexibility subscale, suggesting that accurate monitoring and evaluation of our own cognitive processes co-occur with greater emotional stability and interpersonal attunement.

Perspective-taking (self–other) followed a similar, though slightly weaker, pattern, correlating with the MentS other-related mentalization score and the IPNB-PFCFS subscores. This finding aligns with the notion that inferring the states of others draws on, but does not wholly overlap with, executive control processes.

In contrast, meta-mentalization (other–self) displayed only weak or very weak correlations to most external measures, supporting the view that estimating how well others understand us taps a subtler ability that current Turkish instruments seldom capture. Nevertheless, meta-mentalization scores were negatively related to several borderline personality features (e.g., abandonment fears, feelings of emptiness, and psychotic states), suggesting that perceiving oneself as opaque to others may co-occur with relational insecurity. A small positive link also emerged between perspective-taking and the BPQ Psychotic States subscale, echoing evidence that over-attribution, or "hyper-mentalizing," can accompany transient psychotic-like experiences.

Wu et al. (2022) found that perspective-taking (self–other) correlated positively with psychopathy, whereas metacognition (self–self) and meta-mentalization (other–self) correlated negatively. Our data partially replicated this pattern: metacognition and meta-mentalization were inversely related to both global and secondary-psychopathy scores, whereas perspective-taking showed a small positive association with the primary (interpersonal–affective) psychopathy factor. This convergence supports the proposition that "interpersonal" psychopathic traits may rely on intact, or even enhanced, perspective-taking skills, whereas impulsive–antisocial traits are generally associated with deficient mentalization (Sandvik et al, 2014). Wu et al. (2022) interpreted the positive association between perspective-taking and psychopathy as a potential marker of overconfidence in social-inferential abilities. This interpretation echoes construal-level research showing that psychological distance can foster self-idealization and inflated competence judgments (Griffin et al, 1990; Kivetz & Tyler, 2007). A similar mechanism may operate in our Turkish sample, where higher primary psychopathy was accompanied by slightly elevated claims of perspective-taking accuracy.

At the maladaptive pole, higher BPQ totals and higher secondary-psychopathy scores on the LSRP were associated with lower metacognition in our sample, consistent with research showing that emotional dysregulation and impulsivity compromise introspective accuracy. These findings indicate that the Turkish IMQ captures the core aspects of social cognition and self-regulation while still indexing unique variance, particularly in the meta-mentalization domain.

The weak performance of the other–self subscale in the Turkish context highlights cultural nuances in how individuals perceive being understood by others and limits its immediate interpretability in applied settings. Thus, its use in research and practice should be considered exploratory until further refinement is undertaken.

Clinical Implications

The Turkish IMQ's psychometric soundness supports its potential use in clinical settings. In particular, the measure may serve as a valuable tool in mentalization-based interventions by helping clinicians assess patients' strengths and difficulties across distinct mentalizing dimensions. The strong associations between the self–self and self–other subscales and measures of empathy, prefrontal functioning, and emotion regulation suggest that the Turkish IMQ could inform case formulation

and treatment planning, especially in therapies targeting emotional insight, interpersonal understanding, and affect modulation. Future research may explore the sensitivity of the IMQ to clinical change and its usefulness in evaluating therapeutic progress over time.

However, the other–self subscale demonstrated limited psychometric performance in the Turkish context. Its weak internal consistency, low correlations with other subscales, and inconsistent associations with external measures suggest that this dimension is not yet suitable for clinical application. Further research is necessary to refine the subscale's item content in culturally sensitive ways before it can be confidently applied in therapeutic contexts. Until then, clinicians should interpret other–self scores with caution and avoid using them as standalone indicators in clinical decision-making.

The inclusion of reliable change indices (RCI) and minimally important differences (MID) enhances the clinical interpretability of the Turkish IMQ by providing actionable thresholds for patient monitoring. In practice, clinicians should compute the change score (Δ=follow-up – baseline) for the total or relevant subscale and compare |Δ| to the thresholds in Table 2. If |Δ|≥RCI, the change is statistically reliable (unlikely due to measurement error); if |Δ|≥MID, the change is clinically noticeable/meaningful. Changes can be classified as follows: (i) |Δ|<MID → no meaningful change; (ii) MID≤|Δ|<RCI → possibly meaningful to the patient but not statistically reliable (the clinician should monitor and corroborate); (iii) RCI≤ |Δ|< MID → reliable but small (the clinician should consider incremental adjustment); and (iv) |Δ|≥RCI and≥MID → reliable and clinically meaningful improvement (Δ>0) or worsening (Δ<0). Apply these rules to the total and subscale scores using the corresponding thresholds in Table 2; considering its lower precision, interpret other–self change estimates with added caution and prioritize total, self–self, and self–other for treatment decisions.

Nonetheless, one limitation of the current study concerns the sample's demographic composition. Although the large sample size strengthens the statistical power and generalizability within certain strata, the overrepresentation of university-educated participants (75%) restricts the applicability of the findings to populations with lower educational or socioeconomic backgrounds. Mentalization abilities may be shaped by broader contextual factors, such as access to education, literacy levels, and exposure to psychological vocabulary (Pluck, 2021). Future studies should aim to replicate these findings in more socioeconomically and educationally diverse samples to enhance the Turkish IMQ's ecological validity and generalizability. Another limitation is the exclusive use of a community sample. Although the current study provides robust evidence for the psychometric validity of the Turkish IMQ in a nonclinical population, its performance in clinical populations (e.g., individuals with psychiatric disorders) remains unknown. The capacity of the scale to discriminate between diagnostic groups and its responsiveness to therapeutic change require investigation. Future research should examine the diagnostic utility of the IMQ and its convergent validity with clinically relevant outcomes in structured samples, particularly in disorders marked by impaired mentalization (e.g., borderline personality disorder, psychosis, or autism spectrum disorder).

CONCLUSION

The current findings confirm the Turkish IMQ as a psychometrically sound and clinically promising tool for assessing mentalization in social interaction. While the self–self and self–other subscales demonstrated strong psychometric performance and clinical utility, the other–self subscale requires further refinement and cultural adaptation. The Turkish IMQ can reliably support both research and practice, provided that the other-self subscale is interpreted with caution.

Ethics Committee Approval: The Işık University Institute of Social Sciences Ethics Committee granted approval for this study (date: 16.04.2024, number: 2024/03).

Conflict of Interest: The author has no conflict of interest to declare.

Financial Disclosure: This research was funded by Işık University Scientific Research Projects Coordination Unit (Project No. 23B105). The author gratefully acknowledges Işık University for its support.

Use of AI for Writing Assistance: During the preparation of this work, the author used ChatGPT 5 to edit the language (grammar, syntax, clarity, and readability) of the original draft. After using this tool, the author reviewed and edited the content as needed and take full responsibility for the content of the publication.

Peer-review: Externally peer-reviewed.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.; DSM-IV-TR). American Psychiatric Press.

Arioli, M., Cattaneo, Z., Ricciardi, E., & Canessa, N. (2021). Overlapping and specific neural correlates for empathizing, affective mentalizing, and cognitive mentalizing: A coordinate-based meta-analytic study. *Human Brain Mapping, 42*(14), 4777–4804.

Bhatt, M. A., Lohrenz, T., Camerer, C. F., & Montague, P. R. (2010). Neural signatures of strategic types in a two-person bargaining game. *Proceedings of the National Academy of Sciences of the United States of America, 107*(46), 19720–19725.

Bora, E. (2021). A meta-analysis of theory of mind and "mentalization" in borderline personality disorder: A true neuro-social-cognitive or meta-social-cognitive impairment? *Psychological Medicine, 51*(15), 2541–2551.

Carruthers, P. (1996). Simulation and self-knowledge: A defence of theory-theory. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 22–38). Cambridge University Press.

Ceylan, V., Kose, S., Akin, E. & Turkcapar, M. H. (2017). Normative data and factorial structure of the Turkish version of The Borderline Personality Questionnaire (Turkish BPQ). *Psychiatry and Clinical Psychopharmacology, 27*(2), 143–151.

Dimitrijevic, A., Hanak, N., Dimitrijevic, A. A., & Marjanovic, Z. J. (2018). The Mentalization Scale (MentS): A self-report measure for the assessment of mentalizing capacity. *Journal of Personality Assessment, 100*, 268–280.

Drigas, A., & Mitsea, E. (2020). The 8 pillars of metacognition. *International Journal of Emerging Technologies in Learning (iJET), 15*(21), 162–178.

Engeler, A., and Yargıç, İ. (2004) *Levenson Psikopati Ölçeği (LSRP)'nin Türkçe uyarlamasının klinik olmayan bir örneklem üzerinde güvenirlik ve geçerliği: Pilot çalışmalar* [Poster sunumu]. Türk Psikoloji Derneği, 13. Ulusal Yıllık Kongresi, İstanbul, Türkiye.

Fiedler, K., Ackerman, R., & Scarampi, C. (2019). Metacognition: Monitoring and controlling one's own knowledge, reasoning and decisions. In R. J. Sternberg & J. Funke (Eds.), *The psychology of human thought: An introduction* (pp. 89–111). Heidelberg University Publishing.

Fonagy, P., & Bateman, A. W. (2019). *Handbook of mentalizing in mental health practice*. American Psychiatric Publishing.

Fonagy, P., Gergely, G., & Jurist, E. L. (2002). *Affect regulation, mentalization and the development of the self*. Routledge.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-Functioning Manual, Version 5.0, for Application to Adult Attachment Interviews*. London: University College London.

Gallese, V., Eagle, M. N., & Migone, P. (2007). Intentional attunement: Mirror neurons and the neural underpinnings of interpersonal relations. *Journal of the American Psychoanalytic Association, 55*(1), 131–176.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493–501.

George, C., Kaplan, N., & Main, M. (1996). *Adult Attachment Interview* (3rd ed.). Unpublished manuscript, University of California at Berkeley, USA.

Griffin, D. W., Dunning, D., & Ross, L. (1990). The role of construal processes in overconfident predictions about the self and others. *Journal of Personality and Social Psychology, 59*(6), 1128–1139.

Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind & Language, 7*(1–2), 120–144.

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*(6), 699–709.

Hisli-Şahin, N., & Varlık-Özsoy, E. (2017). *Interpersonal neurobiology-based prefrontal cortex functions scale*. Poster presented at the 8th Işık Savaşır Clinical Psychology Symposium.

Ho, M. K., Saxe, R., & Cushman, F. (2022). Planning with theory of mind. *Trends in Cognitive Sciences, 26*(11), 959–971.

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition, 140*, 14–23.

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 the Mind in the Eyes Test across psychiatric disorders. *Clinical Psychology: Science and Practice, 29*(4), 423–439.

Kivetz, Y., & Tyler, T. R. (2007). Tomorrow I'll be me: The effect of time perspective on the activation of idealistic versus pragmatic selves. *Organizational Behavior and Human Decision Processes, 102*(2), 193–211.

Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., De Roover, K., Jak, S., Meitinger, K., Menold, N., Muthén, B., Rudnev, M., Schmidt, P., & van de Schoot, R. (2023). Measurement invariance in the social sciences: Historical development, methodological challenges, state of the art, and future perspectives. *Social Science Research, 110*, 102805.

Luyten, P., & Fonagy, P. (2018). The neurobiology of attachment and mentalizing: A neurodevelopmental perspective. In C. Schmahl, K. L. Phan, & R. O. Friedel (Eds.), *Neurobiology of personality disorders*. Oxford University Press.

Majdandžić, J., Amashaufer, S., Hummer, A., Windischberger, C., & Lamm, C. (2016). The selfless mind: How prefrontal involvement in mentalizing with similar and dissimilar others shapes empathy and prosocial behavior. *Cognition, 157*, 24–38.

Osterhaus, C., & Bosacki, S. L. (2022). Looking for the lighthouse: A systematic review of advanced theory-of-mind tests beyond preschool. *Developmental Review, 64*, 1–23.

Pluck, G. (2021). Theory of mind ability and socioeconomic status: A study of street-connected children and adolescents in Ecuador. *Psych, 3*(2), 72–84.

Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human metacognition across domains: Insights from individual differences and neuroimaging. *Personality Neuroscience, 1*, e17.

Sandvik, A. M., Hansen, A. L., Johnsen, B. H., & Laberg, J. C. (2014). Psychopathy and the ability to read the "language of the eyes": Divergence in the psychopathy construct. *Scandinavian Journal of Psychology, 55*(6), 585–592.

Silston, B., Bassett, D. S., & Mobbs, D. (2018). How dynamic brain networks tune social behavior in real time. *Current Directions in Psychological Science, 27*(6), 413–421.

Sloover, M., van Est, L. A. C., Janssen, P. G. J., Hilbink, M., & van Ee, E. (2022). A meta-analysis of mentalizing in anxiety disorders, obsessive–compulsive and related disorders, and trauma and stressor related disorders. *Journal of Anxiety Disorders, 92*, 102641.

Stefana, A., Damiani, S., Granziol, U., Provenzani, U., Solmi, M., Youngstrom, E. A., & Fusar-Poli, P. (2025). Psychological, psychiatric, and behavioral sciences measurement scales: Best practice guidelines for their development and validation. *Frontiers in Psychology, 15*, 1494261.

Stefana, A., Jolić Marjanović, Z., & Dimitrijević, A. (2024). The brief version of the Mentalization Scale (MentS-12): Evidence-based assessment of mentalizing capacity. *Journal of Personality Assessment, 106*(6), 740–749.

Taubner, S., White, L. O., Zimmermann, J., Fonagy, P., & Nolte, T. (2013). Attachment-related mentalization moderates the relationship between psychopathic traits and proactive aggression in adolescence. *Journal of Abnormal Child Psychology, 41*(6), 929–938.

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.

Wu, H., Fung, B. J., & Mobbs, D. (2022). Mentalizing during social interaction: The development and validation of the Interactive Mentalizing Questionnaire. *Frontiers in Psychology, 12*, 791835.

Wu, H., Liu, X., Hagan, C. C., & Mobbs, D. (2020). Mentalizing during social interaction: A four component model. *Cortex, 126*, 242–252.

Zhang, S., Howes, A., & Jokinen, J. P. (2025). Human mentalizing as rational probabilistic inference. *Computational Brain & Behavior*, 1–18.

Appendix 1. Descriptive statistics for items in the IMQ (n=953)

Item	Mean (SD)	Skewness	Kurtosis	Min–Max	Item–total correlation
1	3.35 (0.84)	-1.05	0.15	1–4	0.17
2	3.34 (0.70)	-0.81	0.28	1–4	0.41
3	3.19 (0.72)	-0.64	0.26	1–4	0.46
4	3.20 (0.71)	-0.56	-0.01	1–4	0.46
5	2.97 (0.71)	-0.27	-0.19	1–4	0.31
6	2.70 (0.82)	-0.22	-0.45	1–4	0.17
7	2.36 (1.03)	0.14	-1.15	1–4	0.08
8	3.15 (0.75)	-0.57	-0.13	1–4	0.32
9	2.68 (0.70)	-0.28	-0.01	1–4	0.13
10	3.11 (0.74)	-0.53	-0.03	1–4	0.45
11	3.26 (0.71)	-0.71	0.26	1–4	0.47
12	2.24 (0.98)	0.23	-1.02	1–4	0.03
13	2.68 (0.86)	-0.13	-0.65	1–4	0.02
14	3.48 (0.70)	-1.27	1.36	1–4	0.37
15	3.21 (0.72)	-0.61	0.07	1–4	0.35
16	3.30 (0.67)	-0.56	-0.16	1–4	0.48
17	3.02 (0.95)	-0.55	-0.78	1–4	0.17
18	2.37 (1.04)	0.18	-1.13	1–4	-0.00
19	3.08 (0.72)	-0.43	-0.08	1–4	0.28
20	3.35 (0.73)	-0.94	0.51	1–4	0.43

IMQ: Interactive Mentalization Questionnaire; SD: Standard deviation; Min: Minimum; Max: Maximum.

Appendix 2. Descriptive statistics for the confirmatory, exploratory, and full samples

Characteristic	Confirmatory (n=453) n (%)	Exploratory (n=500) n (%)	Full (n=953) n (%)	p
Gender				0.412
Male	265 (59)	280 (56)	545 (57)	
Female	187 (41)	220 (44)	407 (43)	
Other	1 (0)	0 (0)	1 (0)	
Age, years	32.67 (12.25)	30.91 (12.03)	31.75 (12.16)	0.026
Education level				0.967
1	4 (1)	4 (1)	8 (1)	
2	46 (10)	55 (11)	101 (11)	
3	341 (75)	376 (75)	717 (75)	
4	62 (14)	65 (13)	127 (13)	
Socioeconomic status				0.259
1	62 (14)	56 (11)	118 (12)	
2	277 (61)	298 (60)	575 (60)	
3	114 (25)	146 (29)	260 (27)	
Prior psychiatric diagnosis				0.407
Yes	68 (15)	86 (17)	154 (16)	
No	385 (85)	414 (83)	799 (84)	
Psychotropic medication				0.171
Yes	37 (8)	55 (11)	92 (10)	
No	416 (92)	445 (89)	861 (90)	
IMQ total score, mean (SD)	52.77 (5.70)	52.72 (5.80)	52.74 (5.75)	0.893
IMQ_SS	26.13 (3.61)	25.96 (3.53)	26.04 (3.57)	0.463
IMQ_SO	15.67 (2.55)	15.58 (2.53)	15.62 (2.54)	0.585
IMQ_OS	8.62 (1.84)	8.81 (1.72)	8.72 (1.77)	0.101

IMQ: Interactive Mentalization Questionnaire; SD: Standard deviation; Min: Minimum; Max: Maximum.

Appendix 3. Item-level exploratory factor analysis results: The first round

Item	F1	F2	F3
1	0.152	0.596	-0.050
2	0.692	0.102	-0.096
3	0.613	-0.018	0.461
4	0.615	-0.101	0.474
5	0.376	-0.060	0.263
6	0.065	0.681	0.034
7	-0.055	-0.073	0.505
8	0.426	0.020	0.104
9	0.016	0.623	-0.007
10	0.534	-0.024	0.443
11	0.711	0.086	-0.002
12	-0.246	0.271	0.465
13	-0.050	0.283	-0.045
14	0.507	0.049	-0.031
15	0.569	0.002	-0.164
16	0.693	0.057	0.079
17	0.242	-0.032	0.018
18	-0.138	0.292	0.175
19	0.555	-0.117	-0.125
20	0.751	0.024	-0.160

Averaging was performed with the averaging method mean (trim=0) across 72 EFAs, varying the following settings: init_comm, criterion_type, k_promax, P_type, and varimax_type.

Appendix 4. Item-level exploratory factor analysis results–second round

Item	F1	F2	F3
1	0.146	-0.091	0.530
2	0.542	0.167	0.096
3	-0.009	0.882	0.036
4	-0.037	0.944	-0.061
5	0.162	0.343	-0.043
6	-0.083	0.047	0.771
8	0.287	0.210	0.021
9	-0.043	-0.042	0.594
10	0.162	0.571	-0.025
11	0.549	0.217	0.078
14	0.484	0.051	0.042
15	0.661	-0.089	-0.042
16	0.538	0.235	0.054
17	0.280	-0.003	-0.070
19	0.637	-0.054	-0.130
20	0.792	-0.022	-0.012

Averaging was performed with the averaging method mean (trim=0) across 72 EFAs, varying the following settings: init_comm, criterion_type, k_promax, P_type, and varimax_type.

Appendix 5. Etkileşimli Zihinselleştirme Ölçeği

Yönerge: Bu ölçeğin her bir maddesi bir kişinin katılabileceği ya da katılmayabileceği bir ifadeden oluşmaktadır. Her madde için, maddenin söylediklerine ne kadar katıldığınızı veya katılmadığınızı belirtin. Lütfen tüm maddelere yanıt verin; hiçbirini boş bırakmayın. Her ifade için sadece bir yanıt seçin. Lütfen olabildiğince doğru ve dürüst yanıt verin. Her bir maddeyi diğer maddelerden bağımsız olarak düşünüp yanıtlayın. Başka bir deyişle, yanıtlarınızda "tutarlı" olma konusunda endişelenmeyin. Aşağıdaki dört maddeden birini seçin: 1 = çok doğru; 2 = biraz doğru; 3 = biraz yanlış; 4=çok yanlış.

1. Tanımadığınız kişilerin sizin zihninizi başkalarından daha iyi okuyabildiğine inanabiliyor musunuz?
2. Neyi neden yaptığım konusunda isabetli bir içgörüye sahibim.
3. Başka bir kişinin ne düşündüğünü anlamakta iyi olduğuma inanıyorum.
4. Başkalarının ne düşündüğünü anlayabileceğim konusunda kendime güveniyorum.
5. Bir film izlerken karakterin bir sonraki adımda ne yapacağını her zaman tahmin edebilirim.
6. Bazen insanların ne düşündüğümü doğrudan anladıklarını düşünüyorum.
7. Tam tersi bir fikir düşünürken de başka birinin fikrini anlayabilirim.
8. Başkalarının sizin ne düşündüğünüzü tahmin edebileceğinden ne kadar eminsiniz?
9. Arkadaşlarıma kıyasla (ortalama olarak), başkalarının ne düşündüğünü tahmin etme konusunda daha iyiyim.
10. Neyi neden düşündüğüme dair isabetli bir içgörüye sahibim.
11. Başkalarının benimle alay edip etmediğini anlayabilirim.
12. Başarısız olduğumda, tam olarak neden başaramadığımı bilirim.
13. Arkadaşlarıma kıyasla (ortalama olarak), kendi düşünce ve davranışlarım hakkında daha iyi bir içgörüye sahibim.
14. Düşüncelerimi kendime saklamakta iyiyimdir.
15. Yeni bir görevi yerine getirdiğimde onu doğru yaptığımdan eminimdir.
16. Kim olduğumu bilmek konusunda kendime güvenim yüksektir.

Ters puanlanan maddeler: 2, 3, 4, 5, 7, 9, 10, 11, 12, 13, 14, 15, 16.

Alt ölçekler:
- **-ben-öteki:** 3, 4, 5, 7, 9.
- **-ben-ben:** 2, 10, 11, 12, 13, 14, 15, 16.
- **-öteki-ben:** 1, 6, 8.