



FUTURITY
OF SOCIAL SCIENCE

DOI: <https://doi.org/10.57125/FS.2026.06.20.08>

How to cite: Zalova-Nuriyeva, U., Mammadova, R., Nasibova, E., Babazade, Z., Rustamov, E., & Maharova, V. (2026). Revisiting the multidimensional structure of empathy: Psychometric evaluation of the Interpersonal Reactivity Index in a non-Western context. *Futurity of Social Sciences*, 4(2), 152-173. <https://doi.org/10.57125/FS.2026.06.20.08>

Revisiting the Multidimensional Structure of Empathy: Psychometric Evaluation of the Interpersonal Reactivity Index in a Non-Western Context

Ulkar Zalova-Nuriyeva*

PhD Candidate - Head of Personality Laboratory, Psychology Scientific Research Institute, Baku, AZ1000, Azerbaijan, <https://orcid.org/0000-0001-6192-2007>

Rahila Mammadova

Research Assistant, Psychology Scientific Research Institute, Baku, AZ1000, Azerbaijan <https://orcid.org/0009-0006-4050-1944>

Emilya Nasibova

Research Assistant, Psychology Scientific Research Institute, Baku, AZ1000, Azerbaijan <https://orcid.org/0009-0005-7482-5499>

Zhala Babazade

Research Assistant, Psychology Scientific Research Institute, Baku, AZ1000, Azerbaijan, <https://orcid.org/0009-0008-6988-9211>

Elnur Rustamov

*PhD - Chairperson of Psychology Scientific Research Institute, Psychology Scientific Research Institute,
Baku, AZ1000, Azerbaijan*

<https://orcid.org/0000-0002-3241-1707>

Vusala Maharova

Research Assistant, Psychology Scientific Research Institute, Baku, AZ1000, Azerbaijan

<https://orcid.org/0009-0000-6890-5058>

***Correspondence email:** *ulkar.z@psixologiyainstitutu.az*.

Received: February 5, 2026 | **Accepted:** April 24, 2026 | **Published:** June 11, 2026

Abstract. This study aimed to adapt the Interpersonal Reactivity Index (IRI) into Azerbaijani and to evaluate its psychometric properties within a non-Western cultural context, specifically Azerbaijan. A sample of 398 participants completed the IRI along with measures of emotional intelligence, aggression, and self-esteem. Exploratory factor analysis provided initial support for the multidimensional structure of empathy. Internal consistency analyses demonstrated acceptable reliability across the final subscales, with Cronbach's alpha coefficients ranging from .62 to .73. However, confirmatory factor analysis indicated that the original model did not fully meet conventional fit criteria. Following item-level refinement, the final model demonstrated a substantially improved fit ($CFI = .966$, $TLI = .961$, $RMSEA = .068$), supporting the structural validity of the four-factor solution. Item Response Theory (IRT) analyses further confirmed adequate item discrimination across subscales, indicating that the instrument effectively differentiates among individuals at different levels of the latent trait. Correlational findings provided partial support for the nomological network of empathy, with expected associations observed between emotional intelligence and aggression, whereas the relationship with self-esteem revealed a more complex pattern, particularly for the affective components of empathy. Measurement invariance analyses supported configural invariance but failed to establish metric and scalar invariance across gender, suggesting potential group differences in item functioning. Overall, the findings support the reliability and validity of the Azerbaijani version of the IRI and highlight the role of cultural context in shaping the structure and functioning of empathy. This study contributes to cross-cultural psychometric research by demonstrating that while the multidimensional nature of empathy is largely preserved, item-level performance may require cultural adaptation.

Keywords: Empathy, Interpersonal Reactivity Index, Psychometric validation, Emotional intelligence, Aggression.

Introduction

Empathy is widely regarded as a fundamental psychological construct and plays a crucial role in interpersonal functioning, emotional regulation, and social adaptation. Recent research has particularly emphasised that empathy is a multidimensional phenomenon encompassing both cognitive and affective components (Das et al., 2025).

The Interpersonal Reactivity Index (IRI), developed by Davis (1983), remains one of the most commonly used instruments for assessing these dimensions and has been widely adapted across different cultures.

Although numerous global studies have confirmed the cross-cultural applicability of the IRI, findings also indicate that its factor structure may vary across cultures and languages (Siu & Shek, 2005; Yaghoubi Jami & Wind, 2022). For instance, validation studies conducted in Chinese and Persian populations have shown that while the original four-factor model is often supported, certain variations in factor loadings and structure may emerge due to cultural interpretations of the empathy construct (Siu & Shek, 2005; Yaghoubi Jami & Wind, 2022). Similarly, recent validation studies in European and Latin American contexts have highlighted the importance of continuously reevaluating the instrument's psychometric properties when applied to new populations.

Despite the widespread international use of the IRI, research on its adaptation in non-Western and post-Soviet contexts remains limited. In particular, the lack of validated instruments to measure empathy in Azerbaijani restricts both empirical research and applied psychological practice in this area. This gap is especially important given the growing emphasis on socio-emotional competencies in recent years within education, mental health, and organisational development settings (Decety & Jackson, 2004).

The relevance of this study lies in the growing need for culturally appropriate psychological instruments that accurately measure empathy across diverse populations. Empathy has been shown to play a significant role in prosocial behaviour, academic achievement, and psychological well-being; therefore, its reliable assessment is essential for both scientific research and practical interventions (Wang & Lee, 2023). Addressing this gap will contribute to the development of psychological assessment practices in Azerbaijan and support evidence-based interventions in educational and clinical contexts.

The primary aim of this study is to adapt the Interpersonal Reactivity Index (IRI) into the Azerbaijani language and to evaluate its psychometric properties, including reliability and construct validity. In particular, the study seeks to examine whether the original multidimensional structure of empathy is preserved within the Azerbaijani cultural context and whether the instrument demonstrates adequate psychometric performance.

In terms of empirical contribution, this study addresses an important gap in Azerbaijani psychological literature by providing a culturally adapted and psychometrically validated measure of empathy. Unlike previous research, which often relied on unidimensional or non-validated instruments, this study rigorously evaluates the factor validity, reliability, and nomological relationships of the IRI in a non-Western context. Furthermore, by examining the structure of empathy across multiple dimensions, it contributes to cross-cultural validation efforts and enhances the generalizability of empathy models.

From a practical perspective, the availability of a validated Azerbaijani version of the IRI provides a reliable measurement tool for both research and applied settings, including educational, clinical, and organisational contexts. This instrument enables a more precise assessment of socio-emotional competencies and facilitates the development of evidence-based interventions to enhance empathy, improve interpersonal relationships, and promote psychological well-being.

Research Problem

The concept of empathy is a fundamental construct in psychology, particularly in understanding interpersonal relationships, emotional regulation, and social behaviour. Despite its importance, there is still a lack of culturally validated measurement tools in many non-Western contexts, including Azerbaijan. The Interpersonal Reactivity Index (IRI), one of the most widely used instruments for assessing multidimensional empathy, has been extensively validated in Western populations; however, its psychometric properties and structural validity in Azerbaijani samples remain insufficiently examined.

It is necessary to analyse this problem at present because the increasing emphasis on socio-emotional competencies in education, mental health, and organisational psychology requires culturally appropriate and psychometrically sound instruments. Without such tools, both research findings and practical interventions may be limited or biased.

This research is expected to bring several benefits to society. A validated Azerbaijani version of the IRI can improve psychological assessment in clinical, educational, and organisational settings, enabling more accurate identification of empathy-related strengths and difficulties. This, in turn, can support the development of interventions to improve communication, reduce conflict, and enhance social well-being.

From a scientific perspective, the study contributes to cross-cultural psychology by testing whether the multidimensional structure of empathy is consistent across non-Western contexts. It also helps expand the existing literature on measurement adaptation and validation processes.

Although empathy has been widely studied internationally, the field remains insufficiently developed in the Azerbaijani context, particularly regarding standardised measurement tools with established validity and reliability.

Through this research, new insights may be gained regarding the cultural expression of empathy and whether its cognitive and affective components are interpreted similarly across different cultural settings.

Research Focus

The main focus of this study is the psychometric adaptation and validation of the Interpersonal Reactivity Index (IRI) for use in the Azerbaijani cultural context. Specifically, the study concentrates on evaluating whether the multidimensional structure of empathy—comprising cognitive and affective components—can be reliably measured among Azerbaijani participants.

The authors' scientific focus is on examining the validity, reliability, and factor structure of the adapted instrument, ensuring it accurately reflects the construct of empathy in this cultural setting.

Research Aim and Research Questions

The primary aim of this study is to adapt the Interpersonal Reactivity Index (IRI) into Azerbaijani and to evaluate its psychometric properties, including its reliability and validity, within an Azerbaijani sample.

Research Questions

1. Does the original four-factor structure of the Interpersonal Reactivity Index (IRI) replicate within the Azerbaijani cultural context?
2. Does the Azerbaijani version of the IRI demonstrate acceptable reliability across its cognitive and affective empathy dimensions?
3. Do the items of the Azerbaijani IRI demonstrate adequate psychometric performance based on item response theory analyses?
4. Does the Azerbaijani version of the IRI demonstrate acceptable convergent, discriminant, and external validity in relation to emotional intelligence, aggression, and self-esteem?
5. Does the factor structure of the Azerbaijani IRI demonstrate measurement invariance across gender groups?

Literature Review

The university environment represents a critical psychosocial context that shapes the quality of students' social relationships, collaboration in group settings, and academic performance. Contemporary research suggests that the effectiveness of interpersonal interactions in this environment largely depends on socio-emotional competencies, particularly empathy (Zhou et al., 2022; Wang & Lee, 2023). Empathy—the capacity to understand others' emotional and psychological states and to respond appropriately—constitutes a core component of emotional intelligence and plays a fundamental role in the development of social functioning (Decety & Jackson, 2004; Smith et al., 2021).

The student period is characterised as a critical developmental stage marked by intensified personality formation and the restructuring of social roles. During this phase, the development of both cognitive and affective dimensions of empathy serves as a key mechanism that facilitates social adjustment, effective participation in group dynamics, and integration into the academic environment (Miller et al., 2020). Higher levels of empathy are associated with more constructive conflict resolution, increased tolerance, and the development of social responsibility (Kumar & Das, 2024; Rustamov et al., 2023). Moreover, empathic competencies function as protective factors in coping with stress and maintaining psychological well-being (Rodriguez & Singh, 2023; Rustamov et al., 2023).

Empathy is a multidimensional psychological construct encompassing both emotional responsiveness and cognitive perspective-taking processes. In this regard, empathy extends beyond sharing others' emotions to include the capacity to understand their perspectives and to generate appropriate behavioural responses (Decety & Jackson, 2004). Neuropsychological research identifies the mirror neuron system as an underlying mechanism of empathy, as it is activated when individuals observe others' actions or emotional expressions, thereby providing a biological basis for empathic responses (Galese & Goldman, 1998). High levels of empathy are linked to prosocial behaviour, whereas low levels are associated with social withdrawal, increased aggression, and heightened risk of psychopathology (Eisenberg & Miller, 1987).

However, excessively high levels of empathy may also entail certain risks. Sustained and intense empathic engagement has been linked to emotional exhaustion, secondary traumatic stress, and depressive symptoms (Figley, 1995). Therefore, empathy functions adaptively only when balanced by effective regulatory mechanisms; otherwise, it may lead to emotional overload and psychological strain (Decety & Meyer, 2008).

A growing body of literature has examined the relationship between empathy and other psychological constructs. Empathy has been positively associated with self-esteem, contributing to more effective social functioning and self-expression (Ciarrochi et al., 2020). Furthermore, empathy is linked to academic achievement, motivation, and the quality of the learning process (Schunk &

DiBenedetto, 2020). In educational contexts, the presence of empathy in teacher-student relationships enhances academic performance and strengthens social-emotional learning processes (Walden University, 2024; Rustamov et al., 2025).

Despite extensive evidence highlighting the central role of empathy in social relationships and psychological well-being, the precise and reliable measurement of this construct remains a significant methodological challenge. This is particularly relevant in cross-cultural contexts, where the availability of psychometrically validated instruments is essential. In the Azerbaijani context, however, there is a limited availability of multidimensionally and systematically validated tools for assessing empathy.

Materials and Methods

The study was conducted with a sample of 398 individuals aged 18 years and older residing in Azerbaijan. Participants were recruited via convenience sampling, and data were collected through an online survey. The sample consisted predominantly of females (87.94%), with a mean age of 21.81 years (SD = 6.79).

Detailed demographic characteristics, including marital status, education level and economic status, are presented in Table 1.

Table 1

Descriptive statistics of the participants

Variable	Category	n	%
Gender	Female	350	87.94
	Male	48	12.06
Marital Status	Single	348	87.44
	Married	45	11.31
	Divorced	4	1.01
Education Level	Bachelor	357	89.70
	Master	28	7.04
	PhD	13	3.27

Economic Status	Very low	12	3.02
	Low	29	7.29
	Medium	327	82.16
	High	23	5.78
	Very high	7	1.76

Ethics

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki (1975, as revised). Ethical approval was obtained from the Ethics Committee of the Psychology Scientific Research Institute in Baku. All participants provided informed consent prior to participation, and their responses were collected anonymously.

Instrument and Procedure

Interpersonal Reactivity Index (IRI)

Empathy was assessed using the Interpersonal Reactivity Index (IRI; Davis, 1983), a multidimensional self-report measure designed to capture both cognitive and affective components of empathy. The instrument consists of 28 items distributed across four subscales: Perspective Taking (PT), Fantasy (FS), Empathic Concern (EC), and Personal Distress (PD), with each subscale comprising seven items. Responses are recorded on a 5-point Likert scale ranging from 0 ("Does not describe me well") to 4 ("Describes me very well"), with higher scores indicating higher levels of empathy within each dimension.

The IRI has demonstrated adequate psychometric properties across diverse cultural contexts, with internal consistency coefficients typically ranging between .70 and .78 (Pulos et al., 2004). In the present study, internal consistency coefficients for each subscale are reported in the Results section.

Rosenberg Self-Esteem Scale (RSES)

Global self-esteem was measured using the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965), a widely used 10-item instrument assessing overall self-worth. Participants responded on a 4-point Likert scale ranging from 1 ("Strongly disagree") to 4 ("Strongly agree"). The scale includes both positively and negatively worded items, with the latter reverse-coded prior to analysis. Higher scores indicate greater self-esteem.

The RSES has demonstrated strong reliability and validity across cultural contexts, with Cronbach's alpha coefficients typically ranging from .77 to .88. Reliability estimates for the current sample are provided in the Results section. The scale was adapted into Azerbaijani by Aliyev et al. (2025).

Trait Emotional Intelligence Questionnaire – Short Form (TEIQue-SF)

Trait emotional intelligence was assessed using the Trait Emotional Intelligence Questionnaire – Short Form (TEIQue-SF; Petrides & Furnham, 2006). This 30-item self-report measure captures four domains: well-being, self-control, emotionality, and sociability. Responses are provided on a Likert-type scale, with higher scores reflecting greater trait emotional intelligence.

The TEIQue-SF has shown high internal consistency in previous research, typically ranging from .80 to .90. In the present study, reliability coefficients are reported in the Results section.

Buss-Perry Aggression Questionnaire – Short Form (BPAQ-SF)

Aggression was measured using the Buss-Perry Aggression Questionnaire – Short Form (BPAQ-SF; Bryant & Smith, 2001). This 12-item instrument assesses four dimensions of aggression: Physical Aggression, Verbal Aggression, Anger, and Hostility, each represented by three items. Participants rated items on a 5-point Likert scale from 1 ("Extremely uncharacteristic of me") to 5 ("Extremely characteristic of me"), with higher scores indicating greater aggression.

Previous studies have reported acceptable internal consistency for the BPAQ-SF (total scale $\alpha \approx .84$), with subscale reliabilities ranging from .65 to .85. Reliability coefficients for the current sample are presented in the Results section.

Data Analysis

All data were initially coded and screened using IBM SPSS Statistics. This included data cleaning, handling of missing values, and preliminary descriptive analyses. All subsequent statistical analyses were conducted in the R programming environment (R Core Team). Exploratory factor analysis (EFA), confirmatory factor analysis (CFA), item response theory (IRT) modelling, and measurement invariance testing were performed using relevant R packages, including psych, lavaan, and mirt. EFA was used to examine the underlying factor structure of the scale, while CFA with the WLSMV estimator was applied to test the hypothesised four-factor model for ordinal data. Model fit was evaluated using standard indices, including CFI, TLI, RMSEA, and SRMR, following conventional cut-off criteria. IRT analyses were conducted using the Graded Response Model to assess item discrimination and measurement precision across levels of the latent trait. Measurement invariance across gender groups was evaluated sequentially at configural, metric, and scalar levels. All statistical decisions were based on established psychometric guidelines, and significance was evaluated at the conventional alpha level of .05.

Results

Prior to conducting factor analyses, the suitability of the data was assessed using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's Test of Sphericity. The overall KMO value was 0.84, indicating good sampling adequacy according to conventional benchmarks (Kaiser, 1974), and suggesting that the inter-item correlations were sufficiently compact to yield reliable factor solutions. Item-level Measures of Sampling Adequacy (MSA) ranged from 0.63 to 0.91, exceeding the minimum acceptable threshold of 0.60 (Field, 2018), indicating that all items were appropriate for inclusion in the factor analysis.

Bartlett's Test of Sphericity was statistically significant, $\chi^2(378) = 2644.53$, $p < .001$, indicating that the correlation matrix significantly differed from an identity matrix. This result confirms sufficient correlations among variables to justify the application of exploratory factor analysis (Tabachnick & Fidell, 2019).

Taken together, these findings provide strong empirical support for proceeding with factor analytic procedures.

Internal consistency analyses were conducted for all IRI subscales using Cronbach's alpha coefficients. The results indicated acceptable reliability for the Perspective Taking ($\alpha = .70$), Fantasy ($\alpha = .73$), and Personal Distress ($\alpha = .71$) subscales. The Empathic Concern subscale demonstrated comparatively low but still acceptable internal consistency in exploratory and cross-cultural research contexts ($\alpha = .62$). Average inter-item correlations ranged from .29 to .35, indicating acceptable item homogeneity across subscales. Detailed reliability coefficients are presented in Table 2.

Table 2

Reliability Coefficients for the Final IRI Subscales

Subscale	Items (n)	Cronbach's α	95% CI	Mean	SD	Average Inter-item r
Perspective Taking (PT)	5	.70	[.65, .74]	2.50	0.71	.32
Fantasy (FS)	5	.73	[.69, .77]	2.50	0.79	.35
Empathic Concern (EC)	4	.62	[.55, .68]	2.50	0.73	.29

Personal Distress (PD)	5	.71	[.66, .75]	2.10	0.74	.32
------------------------	---	-----	------------	------	------	-----

Note. α = Cronbach's alpha; CI = confidence interval.

These findings suggest that the final retained items demonstrated acceptable internal consistency and were suitable for subsequent factor analytic procedures.

Exploratory Factor Analysis (EFA)

An exploratory factor analysis was conducted to examine the underlying structure of the Azerbaijani version of the Interpersonal Reactivity Index. The model demonstrated an overall acceptable fit to the data (see Table 3).

Table 3

Exploratory Factor Analysis Model Fit Indices

Fit Index	Value
RMSEA	0.044
SRMR	0.040
TLI	0.868
χ^2	484.85
Variance Explained	32%

Specifically, the RMSEA value of 0.044 indicated a close fit, as values below .05 are considered indicative of good model fit (Hu & Bentler, 1999). Similarly, the SRMR value (0.040) suggested minimal discrepancy between observed and reproduced correlations, further supporting model adequacy.

The Tucker-Lewis Index (TLI = 0.868) indicated an acceptable, though not optimal, level of fit. While values above .90 are typically preferred, values above .85 are often considered acceptable in exploratory contexts, particularly in early stages of scale adaptation (Marsh et al., 2004). The chi-square statistic was significant ($\chi^2 = 484.85$, $p < .001$); however, given the known sensitivity of the chi-square test to sample size, this result was interpreted alongside alternative fit indices (Kline, 2016).

The extracted factors accounted for 32% of the total variance, which can be considered moderate but acceptable in psychological measurement, particularly for multidimensional constructs such as empathy (Hair et al., 2019). Overall, the EFA results suggest that the factor structure is structurally meaningful, although further refinement may improve model fit and explanatory power.

Prior to confirmatory factor analysis, item-level diagnostics—including factor loadings and item discrimination indices derived from IRT analyses—were examined to optimise measurement quality. Based on established criteria (e.g., factor loadings < .30, low discrimination, or theoretical inconsistency; Brown, 2015), a total of nine items were removed.

Specifically, two items from the Fantasy subscale (FS7, FS12), two from Personal Distress (PD13, PD19), three from Empathic Concern (EC4, EC14, EC18), and two from Perspective Taking (PT3, PT15) were excluded (see Table 4). This refinement step aimed to improve the scale’s internal consistency and structural validity prior to confirmatory analysis.

Table 4
Items Removed Prior to Confirmatory Factor Analysis

Subscale	Removed Items
Fantasy (FS)	FS7, FS12
Personal Distress (PD)	PD13, PD19
Empathic Concern (EC)	EC4, EC14, EC18
Perspective Taking (PT)	PT3, PT15

A confirmatory factor analysis was conducted on the refined item set to evaluate the hypothesised four-factor structure of the IRI. First, the model was tested using the refined item set, and the initial CFA results indicated a marginal model fit

Table 5
Confirmatory Factor Analysis Fit Indices

Fit Index	Value
χ^2 (df = 146)	338.58***
CFI	0.886
Robust CFI	0.911
TLI	0.866
Robust TLI	0.895
RMSEA	0.058
RMSEA (90% CI)	0.050-0.066
Robust RMSEA	0.050
SRMR	0.059

***p < .001

The Comparative Fit Index (CFI = 0.886) and Tucker-Lewis Index (TLI = 0.866) suggested borderline model fit according to conventional cutoffs ($\geq .90$). Although the robust indices (robust CFI = 0.911; robust TLI = 0.895) indicated improved fit, the overall results suggested that the model did not fully meet optimal fit criteria.

To address this, a second CFA was conducted using the WLSMV estimator, which is recommended for ordinal data and non-normal distributions (Li, 2016).

The refined model demonstrated a good-to-very-good fit to the data (see Table 6).

Table 6

Model Fit Indices for the WLSMV Confirmatory Factor Analysis

Fit Index	Value
-----------	-------

CFI	0.966
TLI	0.961
RMSEA	0.068
SRMR	0.063

Specifically, the Comparative Fit Index (CFI = .966) and Tucker-Lewis Index (TLI = .961) exceeded the recommended threshold of .95, indicating excellent model fit (Hu & Bentler, 1999). The Standardised Root Mean Square Residual (SRMR = .063) was below the acceptable cutoff of .08, suggesting a good fit between observed and model-implied correlations. Although the RMSEA value (.068) slightly exceeded the more stringent criterion of .06, it remained within the acceptable range (< .08), indicating reasonable approximation error (Kline, 2016).

Taken together, these results indicate that the refined model provides a substantially improved representation of the data compared to the initial model and supports the structural validity of the four-factor solution.

Item Response Theory Analysis

To further examine the psychometric properties of the adapted scale at the item level, an Item Response Theory (IRT) analysis was conducted using the Graded Response Model (Samejima, 1969), which is appropriate for ordered categorical response formats such as Likert-type scales. This approach allows for a more precise evaluation of item discrimination and measurement precision across levels of the latent trait (Embretson & Reise, 2000).

The results indicated that all subscales demonstrated acceptable to strong psychometric performance (see Table 7). Item discrimination parameters (a) ranged from 0.49 to 0.75, indicating moderate to high discrimination by conventional benchmarks (Baker, 2001). This indicates that the items were effective in differentiating between individuals with varying levels of empathy.

Among the subscales, the Fantasy (FS) dimension demonstrated the strongest psychometric performance, with the highest mean loading (0.67) and variance explained (45.1%), indicating a well-defined latent structure. The Perspective Taking (PT) and Personal Distress (PD) subscales also showed robust performance, with variance explained exceeding 40%, suggesting coherent measurement of both cognitive and affective aspects of empathy. In contrast, the Empathic Concern (EC) subscale

exhibited comparatively lower indices (variance explained = 39.1%), indicating relatively weaker structural consistency, though still within acceptable limits.

Overall, these findings support the adequacy of item functioning and provide additional evidence of the scale's structural validity at the item level.

Table 7

Item Response Theory Results for the Interpersonal Reactivity Index

Subscale	Items	Range of Loadings (a)	Mean Loading	SS Loadings	Variance Explained
FS (Fantasy)	5	0.51 – 0.75	0.67	2.254	45.1%
PT (Perspective Taking)	5	0.55 – 0.75	0.65	2.112	42.2%
EC (Empathic Concern)	4	0.57 – 0.64	0.61	1.562	39.1%
PD (Personal Distress)	5	0.49 – 0.75	0.63	2.146	42.9%

To assess construct validity, Pearson correlation analyses were conducted between the IRI subscales and theoretically related constructs, including emotional intelligence, aggression, and self-esteem (see Table 8).

Consistent with theoretical expectations, Perspective Taking and Empathic Concern demonstrated positive associations with emotional intelligence, supporting convergent validity (Petrides & Furnham, 2006). Additionally, negative correlations were observed between empathy dimensions and aggression, aligning with prior research suggesting that higher empathy is associated with lower aggressive tendencies (Jolliffe & Farrington, 2006).

Table 8

Correlations Among Study Variables

Variable	1	2	3	4	5	6	7
1. Perspective Taking	1						
2. Fantasy	.30**	1					

3. Empathic Concern	.43**	.43**	1				
4. Personal Distress	.17**	.31**	.32**	1			
5. Emotional Intelligence	.11*	.22**	.17**	.20**	1		
6. Aggression	-.17**	.24**	.04	.20**	.40**	1	
7. Self-esteem	-.13**	-.18**	-.20**	-.24**	.10*	.14**	1

p < .05, ** p < .01

However, the negative correlations observed between empathy and self-esteem warrant careful interpretation. While some studies report positive associations, others suggest that certain components of empathy—particularly Personal Distress—may be linked to lower self-evaluative processes due to heightened emotional sensitivity (Neff, 2003). Thus, these findings may reflect the multidimensional and context-dependent nature of empathy.

Overall, the pattern of correlations provides partial support for the construct's nomological network, indicating acceptable convergent and discriminant validity.

Measurement Invariance Across Gender

Measurement invariance across gender was assessed to determine whether the IRI's factor structure operates equivalently across groups. Configural invariance was supported, indicating that the basic factor structure is consistent across genders (see Table 9). However, metric and scalar invariance were not achieved, as indicated by significant changes in χ^2 and fit indices.

Table 9

Measurement Invariance Across Gender

Model	df	χ^2	$\Delta\chi^2$	Δdf	RMSEA	p
Configural	292	557.62	—	—	—	—
Metric	307	589.02	31.40	15	0.074	.007
Scalar	322	619.77	30.75	15	0.073	.009

The lack of metric and scalar invariance suggests that factor loadings and item intercepts differ across groups, indicating potential gender-related differences in the interpretation or functioning of certain items (Chen, 2007). This finding may be partly attributed to the imbalance in group sizes, particularly the relatively small male subsample.

Given these results, partial invariance approaches or further item refinement may be necessary in future studies.

Discussion

The present study aimed to adapt the Interpersonal Reactivity Index (IRI) into the Azerbaijani context and to evaluate its psychometric properties. Overall, the findings provide substantial support for the multidimensional structure of empathy and demonstrate that the adapted instrument possesses acceptable reliability and validity.

The results of the factor analytic procedures offer important insights into the structural validity of the scale. The exploratory factor analysis (EFA) indicated an acceptable factor structure, supporting the multidimensional nature of empathy. However, the confirmatory factor analysis (CFA) revealed that the initial model did not fully meet conventional fit criteria, necessitating item refinement. Following the removal of poorly performing items, the refined model demonstrated a substantially improved fit, as evidenced by strong CFI and TLI values and acceptable RMSEA and SRMR indices.

This stepwise improvement is consistent with best practices in scale adaptation, where iterative refinement is often required to achieve optimal model fit (Brown, 2015). Importantly, the final model supports the theoretical framework proposed by Davis (1983), confirming that empathy is best conceptualised as a multidimensional construct encompassing both cognitive and affective components. These findings are in line with previous cross-cultural adaptations of the IRI, which have similarly reported the need for item-level modifications to achieve structural adequacy (Pérez-Albéniz et al., 2003; Gilet et al., 2013).

The IRT analysis further strengthened the psychometric evaluation by demonstrating that the items exhibited moderate-to-high discrimination indices. This indicates that the scale effectively differentiates between individuals with varying levels of empathy, a critical requirement for psychological measurement (Embretson & Reise, 2000).

Interestingly, the Fantasy subscale emerged as the most robust dimension, showing the highest explained variance and strongest loadings. This may reflect the cultural salience of imaginative engagement and emotional identification in the Azerbaijani context. In contrast, the relatively lower performance on the Empathic Concern subscale suggests that affective empathy components may be

more context-dependent or culturally influenced, a finding reported in prior cross-cultural empathy research (Chung, 2014).

The pattern of correlations provides partial support for the nomological network of empathy. As expected, empathy dimensions were positively associated with emotional intelligence, supporting the conceptual overlap between these constructs (Petrides & Furnham, 2006). Additionally, the negative relationship between empathy and aggression aligns with previous findings suggesting that empathic individuals are less likely to engage in aggressive behaviour (Jolliffe & Farrington, 2006).

However, the negative associations observed between empathy and self-esteem warrant further discussion. While prior research often reports positive relationships, the present findings may reflect the complexity of empathy, particularly the role of Personal Distress. High levels of affective empathy may increase emotional vulnerability, which in turn may negatively affect self-evaluative processes (Neff, 2003). This interpretation is consistent with theoretical models that distinguish between adaptive and maladaptive components of empathy (Decety & Meyer, 2008).

The measurement invariance analysis indicated that configural invariance was supported, suggesting that the overall factor structure is consistent across gender groups. However, the lack of metric and scalar invariance indicates that certain items may function differently across groups.

This finding should be interpreted with caution, as it may be influenced by an imbalance in sample size, particularly the relatively small number of male participants. Previous studies have shown that gender differences in emotional processing and expression may affect responses to empathy-related items (Christov-Moore et al., 2014). Therefore, future research should aim to test partial invariance models or collect more balanced samples to better understand these differences.

From a theoretical perspective, the findings reinforce the multidimensional conceptualisation of empathy and highlight the importance of distinguishing between cognitive and affective components. The results also contribute to the cross-cultural literature by demonstrating that while the overall structure of empathy is relatively stable, item-level functioning may vary across cultural contexts.

From an applied perspective, the findings have important implications for educational settings, particularly for teacher training, student well-being, and peer relations. The validated Azerbaijani version of the IRI provides a reliable tool for assessing various dimensions of empathy among students and educators, enabling more targeted socio-emotional interventions. For instance, the distinction between cognitive empathy (e.g., Perspective Taking) and affective components (e.g., Empathic Concern and Personal Distress) can inform the design of teacher-training programs aimed at enhancing classroom communication, emotional attunement, and conflict-management skills. Furthermore, the negative association between Personal Distress and self-esteem suggests that interventions should not only promote empathy but also support emotional regulation and self-compassion to prevent empathic

over-arousal and burnout among students. In higher education contexts, the IRI can be used to monitor student well-being and identify individuals at risk of emotional maladjustment. Additionally, the scale may contribute to the development of peer-based programs that foster prosocial behaviour, reduce aggression, and strengthen supportive interpersonal relationships within educational environments. Overall, integrating empathy assessment into educational practice aligns with broader goals of promoting socio-emotional competence and holistic student development.

Limitations

Several limitations should be acknowledged. First, the use of a convenience sampling strategy limits the generalizability of the findings beyond the studied population. Second, the pronounced gender imbalance in the sample may have influenced the results of the measurement invariance analyses, particularly in relation to group comparisons. Third, using a single dataset for both exploratory and confirmatory factor analyses may introduce bias; therefore, future studies are encouraged to employ independent samples to cross-validate the factor structure.

Additionally, the relatively lower performance of the Empathic Concern subscale suggests that further refinement of certain items may be necessary. Cultural factors influencing the expression and interpretation of affective empathy may also play a role and warrant further investigation. Finally, the cross-sectional design limits the ability to draw causal inferences regarding the relationships between empathy and related psychological constructs.

Conclusion

In conclusion, the present study provides empirical support for adapting the Interpersonal Reactivity Index (IRI) in the Azerbaijani context, demonstrating that empathy can be reliably and validly assessed as a multidimensional construct. The integration of classical test theory, confirmatory factor analysis, and item response theory strengthens the methodological rigour of the validation process and highlights the value of a multi-analytic approach in psychometric research.

The findings also underscore the importance of cultural context in shaping item functioning and measurement properties, particularly in relation to affective dimensions of empathy. Despite certain limitations, the adapted Azerbaijani version of the IRI represents a valuable instrument for future research and applied settings, contributing to a more comprehensive understanding of empathy across cultural contexts.

Suggestions for Future Research

Future research should aim to replicate the findings using larger and more gender-balanced samples to ensure the stability of the measurement model across groups. Longitudinal designs are also recommended to examine the temporal stability of empathy and its underlying structure. Moreover,

further studies should investigate cultural moderators that may influence the functioning of affective empathy items, particularly within collectivist or transitional cultural contexts such as Azerbaijan. Experimental and intervention-based studies may also provide valuable insights into how empathy develops and can be enhanced over time.

Acknowledgments

There are no acknowledgements.

Declaration of Conflicting Interests

The authors declare no conflicts of interest.

Funding

The authors received no financial support for the research, authorship, or preparation of this manuscript.

Data Availability Statement

The datasets generated and analysed in the current study are not publicly available but are to be obtained from the corresponding author upon reasonable request.

References

- Aliyev, B., Rustamov, E., Nuriyeva, U. Z., Huseinova, L., Yusifova, T., & Mammadova, R. (2025). Adaptation of the Rosenberg Self-Esteem Scale for use among Azerbaijani adults: Psychometric evaluation and associations with life satisfaction, resilience, depression, anxiety, and stress. *ISRG Journal of Arts, Humanities and Social Sciences*, 3(2). <https://doi.org/10.5281/zenodo.15266368>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.
- Ciarrochi, J., Deane, F. P., & Anderson, S. (2020). Emotional intelligence moderates the relationship between stress and academic performance. *Personality and Individual Differences*, 152, 109584. <https://doi.org/10.1016/j.paid.2019.109584>
- Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: Gender effects in brain and behavior. *Neuroscience & Biobehavioral Reviews*, 46, 604–627. <https://doi.org/10.1016/j.neubiorev.2014.09.001>
- Chung, J. M., Robins, R. W., Trzesniewski, K. H., Nofhle, E. E., & Roberts, B. W. (2014). Continuity and change in self-esteem during emerging adulthood. *Journal of Personality and Social Psychology*, 106(3), 469–483. <https://doi.org/10.1037/a0035135>

- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3(2), 71–100. <https://doi.org/10.1177/1534582304267187>
- Decety, J., & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and Psychopathology*, 20(4), 1053–1080. <https://doi.org/10.1017/S0954579408000503>
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113–126. <https://doi.org/10.1037/0022-3514.44.1.113>
- Das, A. K., Nayak, J., Naik, B., Maringanti, H. B., Vimal, S., & Pelusi, D. (Eds.). (2025). *Computational intelligence in pattern recognition: Proceedings of CIPR 2024* (Vol. 2). Springer.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates Publishers.
- Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial behavior. *Psychological Bulletin*, 101(1), 91–119. <https://doi.org/10.1037/0033-2909.101.1.91>
- Figley, C. R. (1995). *Compassion fatigue*. Brunner/Mazel.
- Galese, V., & Goldman, A. (1998). Mirror neurons and simulation theory. *Trends in Cognitive Sciences*, 2(12), 493–501. [https://doi.org/10.1016/S1364-6613\(98\)01262-5](https://doi.org/10.1016/S1364-6613(98)01262-5)
- Jolliffe, D., & Farrington, D. P. (2006). Basic Empathy Scale. *Journal of Adolescence*, 29(4), 589–611. <https://doi.org/10.1016/j.adolescence.2005.08.010>
- Miller, T., Smith, A., & Johnson, L. (2020). Cognitive and affective empathy. *Journal of Youth Studies*, 23(4), 452–468. <https://doi.org/10.1080/13676261.2019.1655610>
- Neff, K. D. (2003). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity*, 2(2), 85–101. <https://doi.org/10.1080/15298860309032>
- Pérez-Albéniz, A., et al. (2003). Spanish adaptation of IRI. *Psicothema*, 15(2), 267–272.
- Petrides, K. V., & Furnham, A. (2006). The role of trait emotional intelligence in a gender-specific model of organizational variables. *Journal of Applied Social Psychology*, 36(2), 552–569. <https://doi.org/10.1111/j.0021-9029.2006.00019.x>
- Rodriguez, M., & Singh, P. (2023a). Empathy and mental health. *Journal of Adolescence*, 99, 345–359. <https://doi.org/10.1016/j.adolescence.2022.12.008>

- Rustamov, E., Musayeva, T., Xalilova, X., et al. (2023). Exploring the relationship between social connectedness and mental wellbeing: The mediating role of psychological resilience among adults in Azerbaijan. *Discover Psychology*, 3, 15. <https://doi.org/10.1007/s44202-023-00080-8>
- Rustamov, E., Zalova Nuriyeva, U., alahverdiyeva, M., Abbasov, T., Mammadzada, G., & Rustamova, N. (2023). Academic self-efficacy, academic procrastination, and well-being: A mediation model with large sample of Azerbaijan. *International Online Journal of Primary Education*, 12(2), 84-93. <https://doi.org/10.55020/iojpe.1250574>
- Rustamov, E., Zalova Nuriyeva, U., & Mammadova, R. (2025). An Azerbaijani adaptation of the Interpersonal Support Evaluation List: Exploring the relationship between interpersonal support, wellbeing, psychological resilience, depression, anxiety and stress. *International Journal of Human Research and Social Science Studies*, 2(5), 276-283. <https://doi.org/10.55677/ijhrsss/07-2025-Vol02I5>
- Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and SEL. *Contemporary Educational Psychology*, 60, 101830. <https://doi.org/10.1016/j.cedpsych.2020.101830>
- Siu, A. M. H., & Shek, D. T. L. (2005). Validation of the Interpersonal Reactivity Index in a Chinese context. *Research on Social Work Practice*, 15(2), 118-126. <https://doi.org/10.1177/1049731504270384>
- Walden University. (2024). Empathy in academic performance. *Journal of Educational Research and Practice*, 12(3), 101-112.
- Wang, Y., & Lee, S. (2023). Empathy and social support. *Journal of College Student Development*, 64(3), 345-359. <https://doi.org/10.1353/csd.2023.0034>
- Yaghoubi Jami, P., & Wind, S. A. (2022). Evaluating the psychometric properties of a proposed Farsi version of the Interpersonal Reactivity Index using item response theory. *Research on Social Work Practice*, 32(8), 1003-1018. <https://doi.org/10.1177/10497315221089322>
- Zhou, L., et al. (2022). Social skills and empathy. *Frontiers in Psychology*, 13, 870123. <https://doi.org/10.3389/fpsyg.2022.870123>