TPM Vol. 32, No. 4, 2025
ISSN: 1972-6325
https://www.tpmap.org/

CROSS-CULTURAL ADAPTATION AND PSYCHOMETRIC EVALUATION OF THE INTERPERSONAL REACTIVITY INDEX (IRI) FOR INDONESIAN PRE-SERVICE EARLY CHILDHOOD TEACHERS

LUTFATULATIFAH
UNIVERSITAS NEGERI MALANG, INDONESIA
UNVERSITAS ISLAM NEGERI SIBER SYEKH NURJATI CIREBON, INDONESIA

HENNY INDRESWARI
UNIVERSITAS NEGERI MALANG, INDONESIA

NUR EVA
UNIVERSITAS NEGERI MALANG, INDONESIA

YULIATI HOTIFAH
UNIVERSITAS NEGERI MALANG, INDONESIA

DONNA D. DE GUZMAN
[3]NUEVA ECIJA UNIVERSITY OF SCIENCE AND TECHNOLOGY, PHILIPPINES

Abstract: Empathy is a vital competence for early childhood teachers, influencing teacher– child interactions and emotional learning. However, the conceptualization and expression of empathy may differ across cultural contexts, making cross-cultural validation essential to ensure that the construct is measured equivalently. This study aimed to adapt and exam- ine the psychometric properties of the Interpersonal Reactivity Index (IRI; Davis, 1983) for Indonesian pre-service early childhood teachers, ensuring its cultural and linguistic rele- vance within the Indonesian educational context. Using Beaton et al.'s (2000) cross-cultural adaptation procedure, the process included translation, synthesis, back-translation, expert review, and pilot testing. Data from 251 participants were analyzed through item validity, exploratory factor analysis (EFA), and confirmatory factor analysis (CFA). Results sup- ported a four-factor structure—Fantasy, Empathic Concern, Perspective Taking, and Per- sonal Distress—explaining 73.25% of the total variance. Model fit indices were acceptable (χ^2 = 791.60; GFI = 0.914; CFI = 0.940; TLI = 0.919; RMSEA = 0.086). Cronbach's α =.976 and composite reliability = 0.87–0.94 confirmed strong internal consistency. The adapted IRI demonstrated satisfactory validity, reliability, and cross-cultural equivalence, providing a psychometrically sound tool for assessing empathy in Indonesian early child- hood teacher education.

Keywords: Confirmatory Factor Analysis; Cross-Cultural Validation; Empathy; Interpersonal Reactivity Index; Psychometric Evaluation.

INTRODUCTION

Empathy is a multidimensional psychological construct essential to effective teaching, especially in early childhood education (Denham et al., 2014; Hoffman, 2000). It enables teachers to understand children's emotions, respond sensitively to behavioral cues, and foster emotionally secure learning environments. For pre-service teachers, empathy functions not only as an interpersonal skill but as a professional disposition influencing classroom management, moral development, and inclusive practices (Indreswari et al., 2022; Jiang et al., 2025; Setyawan et al., 2021). Within early childhood settings, where learning is inherently relational, empathy supports teachers in scaffolding emotional regulation, recognizing individual needs, and cultivating caring classroom climates.

The Interpersonal Reactivity Index (IRI) developed by (Davis, 1983) is one of the most widely used self-report measures for empathy, consisting of four subscales: Perspective Taking (PT), Fantasy (FS), Empathic Concern (EC), and Personal Distress (PD). Despite its robustness, empirical applications of the IRI in teacher-education contexts remain limited, particularly in non-Western cultures. Cultural psychology emphasizes that empathy is both universal and culturally shaped while its basic emotional mechanisms may be similar across humanity, its expression, interpretation, and salience vary across societies (Atkins et al., 2016; Pedersen, 2009). In individualist cultures, empathy is often conceptualized as understanding others' perspectives as separate and autonomous individuals. In collectivist cultures, such as Indonesia, empathy is expressed more relationally, emphasizing social harmony, mutual adjustment, and the maintenance of group

cohesion. Consequently, the meaning and measurement of empathy require cultural contextualization to ensure that instruments capture the construct as it is actually experienced and valued within specific sociocultural frameworks. Cultural and linguistic differences may alter how empathy is expressed and interpreted, necessitating contextual validation of the instrument (Beaton et al., 2000; Hambleton & de Jong, 2003).

In Indonesia, early childhood teachers are expected to embody relational sensitivity, emotional attunement, and social care traits that align closely with the cultural values of gotong royong and kebersamaan. These expectations position empathy as a foundational component of pedagogical practice and professional ethics. Pre-service early childhood teachers represent a particularly important population for empathy measurement because they are in the formative stage of developing professional identity and emotional competence. Their empathy levels influence how they will later interpret children's emotions, manage behavior, and respond to diversity and inclusion in classrooms. Assessing empathy among pre-service teachers therefore provides critical insight into the social-emotional dimensions of teacher preparation and highlights areas for targeted intervention in teacher education curricula.

From a theoretical perspective, empathy encompasses both cognitive and affective dimensions (Davis, 1983; Hoffman, 2000). The cognitive dimension, represented by Perspective Taking and Fantasy, involves the ability to understand another's internal state and imaginatively adopt their viewpoint. The affective dimension, reflected in Empathic Concern and Personal Distress, pertains to emotional responsiveness and affective resonance with others' experiences. In early childhood education, the integration of these two dimensions underpins teacher–child interactions that are emotionally responsive and developmentally supportive. Teachers who can accurately perceive children's emotional cues (cognitive empathy) and respond with genuine warmth and concern (affective empathy) are more likely to create secure, trusting relationships that facilitate learning and emotional growth (Denham et al., 2014).

Therefore, adapting and validating the IRI in Bahasa Indonesia is a necessary step toward ensuring that empathy can be measured in ways that reflect Indonesian cultural norms and teacher-education contexts(Abidin et al., 2020; Hotifah et al., 2022; Supriatna et al., 2024). This study aimed to (1) translate and culturally adapt the Interpersonal Reactivity Index (IRI) for Indonesian pre-service early childhood teachers and (2) examine its psychometric properties through item analysis, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and reliability testing.

METHOD

Design

This study employed a methodological design aimed at instrument adaptation and validation (instrumental adaptation study), following the cross-cultural adaptation guidelines proposed by (Beaton et al., 2000). The process involved translation, cultural adaptation, pilot testing, and psychometric evaluation through both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA).

Adaptation Procedure

The adaptation was conducted through five sequential stages:

1. T1–T2 translation. Two bilingual translators independently translated the original English version of the Interpersonal Reactivity Index (IRI) (Davis, 1983) into Bahasa Indonesia, one with an educational background and the other from a non-educational discipline to ensure linguistic and contextual breadth.
2. Synthesis (T12 version). The two translations were compared and synthesized into a unified version after discussion of linguistic discrepancies.
3. Back translation. The synthesized version was translated back into English by a sworn translator and verified by the Balai Bahasa, ensuring conceptual equivalence with the source instrument.
4. Expert committee review. A panel of five experts—representing linguistics, education, and early-childhood education—evaluated the pre-final Indonesian version for semantic, idiomatic, experiential, and conceptual equivalence. Each expert independently rated the relevance and clarity of each item using a 4-point scale. The results showed Aiken's V coefficients ranging from 0.80 to 1.00 (M = 0.94), indicating excellent inter-rater agreement and confirming that all 28 items achieved satisfactory content validity.
5. Pilot testing. The pre-final version was administered to 40 pre-service early-childhood teachers to evaluate clarity and comprehensibility prior to large-scale data collection.

Psychometric Evaluation

Following the adaptation process, the psychometric evaluation was conducted to examine the construct validity and reliability of the Indonesian version of the Interpersonal Reactivity Index (IRI). The evaluation consisted of several stages:

1. Item analysis and internal reliability (Cronbach's α, Composite Reliability).
2. Exploratory Factor Analysis (EFA) to identify the underlying factor structure.
3. Confirmatory Factor Analysis (CFA) to verify the model fit and factorial validity using AMOS v.26.
4. Convergent validity using Average Variance Extracted (AVE).
5. Model fit evaluation using CFI, TLI, RMSEA, and SRMR indices.

Participants

The main study involved 251 students enrolled in Early Childhood Teacher Education programs (PGPAUD/PIAUD) from several teacher-training institutes (LPTK) across Indonesia. Participants ranged in age from 18 to 23 years, reflecting the typical demographic of pre-service early childhood teachers. The sample was predominantly female (97.4%), with male participants comprising 2.6% of the total. Regarding teaching experience, 40% of the respondents had no prior teaching experience, while 60% reported having taught in early childhood settings either during practicum or as part-time assistants. Participation in the study was voluntary and anonymous.

Instrument

The adapted instrument comprised 28 items from the original Interpersonal Reactivity Index (IRI; Davis, 1983), measured on a 4-point Likert scale ranging from 1 (does not describe me at all) to 4 (describes me very well). The scale retained the original four theoretical dimensions—Fantasy (FS), Empathic Concern (EC), Perspective Taking (PT), and Personal Distress (PD)—which represent cognitive and affective components of empathy.

A total of nine items were negatively worded (unfavorable) to control for acquiescence bias and were reverse-coded prior to data analysis. Minor lexical adaptations were made to ensure cultural relevance and contextual fit for the early-childhood-education setting (e.g., wording adjustments to reflect teacher–child interactions and emotional understanding within the classroom context).

Data Analysis

Data were analyzed using both classical and structural-equation approaches.

1. Item validity was assessed via Pearson correlations between item and total scores ($r = .64–.89$, $p < .001$).
2. Internal reliability was tested through Cronbach's $\alpha = .976$ and Composite Reliability (CR = 0.87–0.94).
3. Exploratory Factor Analysis (EFA) (Principal Component with Varimax rotation) verified the underlying structure prior to CFA. Sampling adequacy was excellent (KMO = .925; Bartlett's $\chi^2 = 8606.36$, df = 378, $p < .001$). Four factors emerged, explaining 73.25% of the total variance.
4. Confirmatory Factor Analysis (CFA) was performed using AMOS v.26 with the maximum-likelihood estimation method to test the four-factor model. Model fit was evaluated using multiple indices: CFI = .940, TLI = .919, GFI = .914, RMSEA = .086, SRMR = .032, indicating an acceptable model fit.
5. Convergent validity was determined through Average Variance Extracted (AVE = 0.53–0.66), exceeding the .50 criterion (Fornell & Larcker, 1981).

Overall, the analytic sequence (EFA → CFA) confirmed that the Indonesian adaptation of the IRI possessed a robust factorial structure, strong internal consistency, and satisfactory construct validity.

RESULTS

Content Validity (Aiken's V)

Five experts assessed semantic(S), idiomatic(I), experiential(E), and conceptual (C) equivalence for all 28 items of the Indonesian IRI. The Aiken's V coefficients ranged from 0.80 to 1.00, with an average of 0.94, indicating very high expert agreement and satisfactory content validity for all items. No items required elimination, and only minor lexical adjustments were made to enhance clarity for early-childhood contexts.

TABLE 1 Aiken's V

Item	Assesed	Expert					Aikens V	
		1	2	3	4	5		
Item 1	S	3	4	4	4	3	0.866667	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	3	4	4	0.933333	High (Retain)
	C	4	4	4	4	3	0.933333	High (Retain)
Item 2	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	3	0.933333	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 3	S	3	4	4	4	3	0.866667	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 4	S	3	4	4	4	4	0.933333	High (Retain)
	I	3	4	3	4	3	0.8	High (Retain)
	E	4	4	3	4	4	0.933333	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 5	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	4	1	High (Retain)

	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 6	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 7	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	3	0.933333	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 8	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	3	0.933333	High (Retain)
	E	4	4	4	4	3	0.933333	High (Retain)
	C	4	4	4	4	3	0.933333	High (Retain)
Item 9	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 10	S	3	4	4	4	4	0.933333	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	3	0.933333	High (Retain)
	C	4	4	4	4	3	0.933333	High (Retain)
Item 11	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 12	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	3	0.933333	High (Retain)
	E	4	4	3	4	3	0.866667	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 13	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	3	0.933333	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	3	0.933333	High (Retain)
Item 14	S	4	4	4	4	4	1	High (Retain)
	I	4	4	3	4	4	0.933333	High (Retain)
	E	4	4	3	4	4	0.933333	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 15	S	4	4	4	3	4	0.933333	High (Retain)
	I	4	4	3	4	4	0.933333	High (Retain)
	E	4	4	3	4	4	0.933333	High (Retain)
	C	4	4	4	4	3	0.933333	High (Retain)
Item 16	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	3	4	4	0.933333	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 17	S	4	4	4	3	4	0.933333	High (Retain)
	I	4	4	3	4	4	0.933333	High (Retain)
	E	4	4	3	4	4	0.933333	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 18	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 19	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 20	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	3	0.933333	High (Retain)

	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 21	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 22	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 23	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 24	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 25	S	3	4	4	4	3	0.866667	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 26	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 27	S	4	4	4	4	3	0.933333	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
Item 28	S	4	4	4	4	4	1	High (Retain)
	I	4	4	4	4	4	1	High (Retain)
	E	4	4	4	4	4	1	High (Retain)
	C	4	4	4	4	4	1	High (Retain)
		442	448	437	446	420		

Item Validity and Reliability

All items exhibited significant item–total correlations, ranging from r = 0.641 to 0.888 ($p < .001$), demonstrating strong internal consistency. The overall Cronbach's α was 0.976, suggesting excellent reliability. Dimension-specific reliability was also satisfactory, with α coefficients exceeding 0.85 across all four empathy subscales (Fantasy, Empathic Concern, Perspective Taking, and Personal Distress).

TABLE 2 Validity Item

Correlations			
		total	
FS1	Pearson Correlation	.839**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
FS2	Pearson Correlation	.878**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
FS3	Pearson Correlation	.813**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
FS4	Pearson Correlation	.731**	Valid

	Sig. (2-tailed)	0.000	
	N	251	
FS5	Pearson Correlation	.835**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
FS6	Pearson Correlation	.750**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
FS7	Pearson Correlation	.769**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
FS8	Pearson Correlation	.825**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
EC1	Pearson Correlation	.696**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
EC2	Pearson Correlation	.801**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
EC3	Pearson Correlation	.888**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
EC4	Pearson Correlation	.699**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
EC5	Pearson Correlation	.641**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
EC6	Pearson Correlation	.850**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
EC7	Pearson Correlation	.746**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PT1	Pearson Correlation	.815**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PT2	Pearson Correlation	.705**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PT3	Pearson Correlation	.690**	Valid

	Sig. (2-tailed)	0.000	
	N	251	
PT4	Pearson Correlation	.792**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PT5	Pearson Correlation	.772**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PT6	Pearson Correlation	.719**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PD1	Pearson Correlation	.697**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PD2	Pearson Correlation	.791**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PD3	Pearson Correlation	.843**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PD4	Pearson Correlation	.847**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PD5	Pearson Correlation	.869**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PD6	Pearson Correlation	.843**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
PD7	Pearson Correlation	.785**	Valid
	Sig. (2-tailed)	0.000	
	N	251	
total	Pearson Correlation	1	
	Sig. (2-tailed)		
	N	251	

TABLE 3 Reliability Item

Reliability Statistics	
Cronbach's Alpha	N of Items
.976	28

N = 251. All correlations significant at $p < .001$. The reliability coefficients exceeded the recommended threshold ($\alpha \geq .70$), indicating excellent internal consistency across all four dimensions of empathy.

Exploratory Factor Analysis (EFA)

Sampling adequacy was confirmed by the Kaiser–Meyer–Olkin (KMO) = 0.925 and Bartlett's Test of Sphericity $\chi^2(378) = 8606.36$, $p < .001$. EFA with Varimax rotation revealed four dominant factors (eigenvalues > 1), explaining 73.25% of total variance. Each factor corresponded to the theoretical dimensions proposed by (Davis, 1983):

1. Fantasy (FS): loadings 0.512–0.774
2. Empathic Concern (EC): loadings 0.505–0.754
3. Perspective Taking (PT): loadings 0.569–0.749
4. Personal Distress (PD): loadings 0.647–0.783

The rotated solution showed a clear, interpretable structure without cross-loading issues, confirming satisfactory construct validity.

TABLE 4

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.925
Bartlett's Test of Sphericity	Approx. Chi-Square	8606.357
	df	378
	Sig.	.000

Before performing the Exploratory Factor Analysis (EFA), sample adequacy was first tested using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's Test of Sphericity. The results are presented in Table 1. The KMO value of 0.925 indicates that the sample size in this study was in the excellent category for factor analysis, as it exceeded the minimum threshold of 0.50. Furthermore, Bartlett's Test yielded a Chi-Square value of 8606.357 with df = 378 and $p < 0.000$, demonstrating that the inter-item correlations were significant and that the correlation matrix was not an identity matrix. Therefore, the data met the assumptions required to proceed with the exploratory factor analysis.

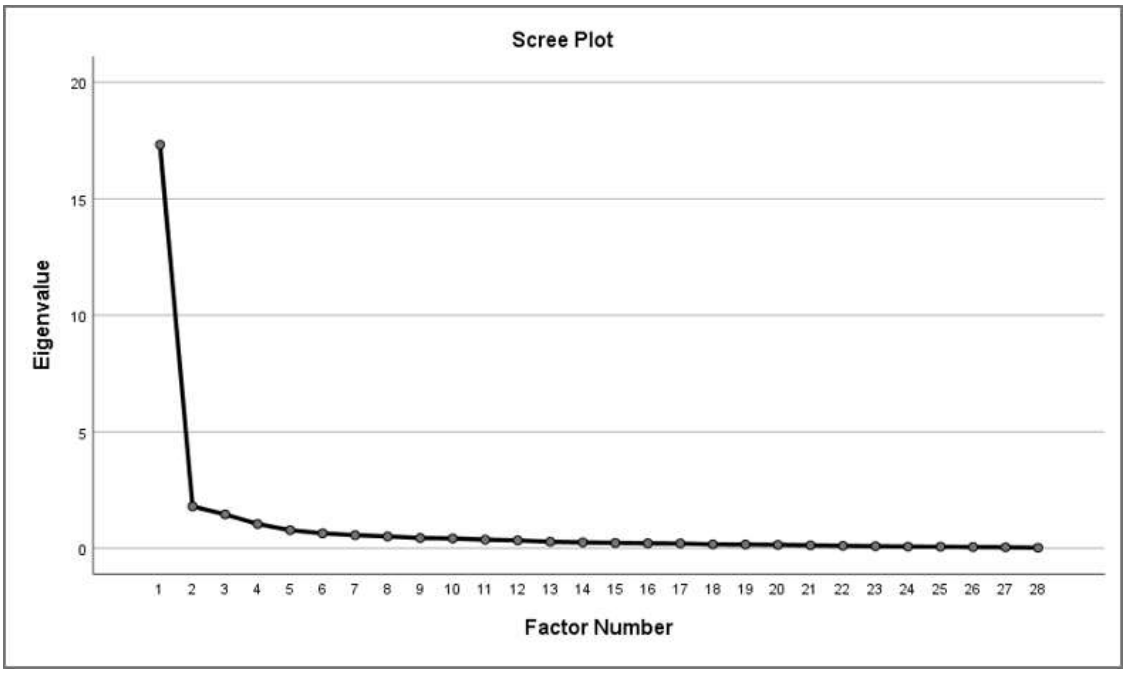
TABLE 5Total variance explained (rotation sums of squared loadings)

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	17.320	61.859	61.859	16.996	60.698	60.698	7.922	28.292	28.292
2	1.801	6.431	68.290	1.391	4.969	65.668	5.511	19.681	47.973
3	1.458	5.206	73.496	1.177	4.202	69.869	4.679	16.709	64.682
4	1.046	3.735	77.231	.946	3.378	73.247	2.398	8.566	73.247

Based on the results of the Exploratory Factor Analysis (EFA) presented in Table 2, four main factors emerged according to the criterion of eigenvalues greater than one. The values in the Initial Eigenvalues column show that the first factor had an eigenvalue of 17.320, explaining 61.86% of the total variance. The second factor had an eigenvalue of 1.801 contributing 6.43%, the third 1.458 contributing 5.21%, and the fourth 1.046 contributing 3.74%. Cumulatively, the four factors explained 77.23% of the total variance before extraction. After the Extraction Sums of Squared Loadings procedure, the four retained factors remained consistent. The cumulative variance explained by these four factors was 73.25%, indicating that the factor model demonstrated a satisfactory level of data explanation.

Furthermore, the Rotation Sums of Squared Loadings column showed that after Varimax rotation, the variance was more evenly distributed and easier to interpret. The first factor explained 28.29% of the variance, the second 19.68%, the third 16.71%, and the fourth 8.56%, with a total cumulative variance of 73.25%. The more balanced variance distribution after rotation indicates that the rotation successfully clarified the factor structure and reduced the dominance of any single factor.

These results suggest that the four extracted factors have strong and adequate explanatory power, as the cumulative variance explained exceeded the commonly accepted minimum threshold of 60% for factor model adequacy in social and psychological research (Tabachnick, 2007). Therefore, it can be concluded that the factor model derived from this analysis meets the required criteria, and the four factors identified are statistically valid representations of the constructs measured in this study.



Picture 1 Scree Plot

The figure above presents the Scree Plot illustrating the relationship between eigenvalues and the number of factors extracted from the Exploratory Factor Analysis (EFA). The plot shows a steep decline in eigenvalues from the first to the second factor, followed by a gradual leveling off after the fourth factor. The visible elbow point at the fourth factor indicates the optimal number of factors to retain in the model. This is consistent with Kaiser's Rule, which states that only factors with eigenvalues greater than 1 should be retained(Cliff, 1988; Ledesma et al., 2021). Accordingly, the Scree Plot confirms the findings in Table 2, indicating that four main factors are appropriate for retention, as the eigenvalues drop and stabilize after the fourth factor. This interpretation strengthens the conclusion that the resulting factor structure meets the adequacy criteria and supports the construct validity of the adapted instrument.

TABLE 6 Rotated Component Matrix (RCM)

Rotated Factor Matrix[a]				
Indikator/ Item	Factor			
	FS	EC	PT	PD
FS1	0.630			
FS2	0.697			
FS3	0.774			
FS4	0.578			
FS5	0.512			
FS6	0.604			
FS7	0.733			
FS8	0.632			
EC1		0.505		
EC2		0.730		
EC3		0.758		
EC4		0.684		
EC5		0.688		

EC6		0.710		
EC7		0.664		
PT1			0.736	
PT2			0.569	
PT3			0.579	
PT4			0.697	
PT5			0.612	
PT6			0.749	
PD1				0.653
PD2				0.660
PD3				0.711
PD4				0.783
PD5				0.738
PD6				0.725
PD7				0.647

The Rotated Component Matrix (RCM) presented in the figure above displays the distribution of factor loadings for each item across the four extracted factors. Factor loadings indicate the strength of the relationship between an item and a given factor; the higher the loading value (ideally ≥ 0.40), the greater the item's contribution to that factor. Based on the rotation results, the four factors show a clear and distinct structure without significant overlap. Each factor contains a cluster of items with dominant loadings on a single column, indicating good construct validity for the model. The interpretation of each factor is as follows:

1. Factor 1 (FS)

This factor consists of items with loading values ranging from 0.512 to 0.774. All items load strongly on the same factor, which can be interpreted as the Fantasy (FS) dimension, reflecting the individual's imaginative capacity and tendency to immerse themselves in hypothetical scenarios.

2. Factor 2 (EC)

The second factor shows loadings between 0.505 and 0.754. These items cluster together and represent the Empathic Concern (EC) dimension, which reflects the individual's affective responsiveness and emotional sensitivity toward others.

3. Factor 3 (PT)

This factor includes items with loadings ranging from 0.569 to 0.746. It corresponds to the Perspective Taking (PT) dimension, which captures the individual's cognitive ability to adopt another person's point of view.

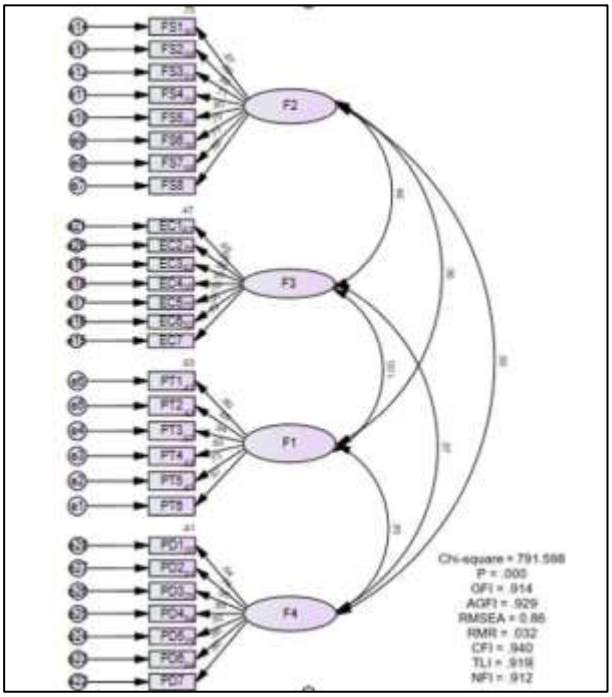
4. Factor 4 (PD)

The fourth factor has loadings between 0.647 and 0.783. This factor represents the Personal Distress (PD) dimension, indicating the individual's tendency to experience discomfort and anxiety in emotionally charged situations.

Overall, the rotated solution demonstrates that each item loads strongly on its designated factor without meaningful cross-loadings, indicating a stable and valid factor structure. These findings are consistent with the Total Variance Explained and Scree Plot results, where all four factors had eigenvalues greater than 1 and cumulatively explained 73.25% of the total variance. Thus, the four extracted factors can be accepted as valid representations of the underlying constructs measured in this study.

Confirmatory Factor Analysis (CFA)

The four-factor model was tested using maximum-likelihood estimation in AMOS (v.26). The model achieved an acceptable level of fit: $\chi^2(320) = 791.60$, $p < .001$, GFI = 0.914, AGFI = 0.929, CFI = 0.940, TLI = 0.919, RMSEA = 0.086, and RMR = 0.032. Although χ^2 was significant—likely due to the large sample size—the comparative and incremental fit indices all exceeded the 0.90 criterion, indicating a good fit between the data and the hypothesized model. The results supported the four-dimensional structure of empathy (FS, EC, PT, PD) in the Indonesian sample.



Picture 2. Confirmatory Factor Analysis (CFA)

Based on the results of the Structural Equation Modeling (SEM) analysis using AMOS, several Goodness-of-Fit indices were obtained as follows: $\chi^2 = 791.598$, p = 0.000, GFI = 0.914, AGFI = 0.929, RMSEA = 0.086, RMR = 0.032, CFI = 0.940, TLI = 0.919, and NFI = 0.912. The chi-square value was significant ($p < 0.05$), indicating that the model did not fully fit the data statistically. However, because chi-square is highly sensitive to large sample sizes ($N > 200$), the evaluation of model fit relies more on alternative fit indices. The results show that most indices meet or closely approach the criteria for a good fit. The GFI (0.914) and AGFI (0.929) values exceed the minimum threshold of 0.90, indicating good model fit. The CFI (0.940), TLI (0.919), and NFI (0.912) values are also above 0.90, suggesting adequate incremental fit. The RMR value (0.032) remains below 0.05, indicating minimal discrepancies between the observed and predicted covariances. Meanwhile, the RMSEA value (0.086) is slightly above the ideal cut-off (≤ 0.08), but still falls within the acceptable range categorized as marginal fit (Hotifah et al., 2022). Overall, the model can be considered to exhibit an acceptable level of fit and is suitable for further analysis of the latent variable relationships in this study.

Convergent Validity and Composite Reliability

Convergent validity was confirmed through high factor loadings (0.566–0.922) and Average Variance Extracted (AVE) values above 0.50, specifically: FS = 0.66, EC = 0.55, PT = 0.53, PD = 0.55. Composite Reliability (CR) was strong across all dimensions: FS = 0.94, EC = 0.89, PT = 0.87, and PD = 0.89, exceeding the recommended threshold of 0.70. These results demonstrate that each construct captures a substantial proportion of its indicator variance and possesses excellent internal consistency.

TABLE 4 Composite Reliability (CR) and Average Variance Extracted (AVE)

Dimension	CR	AVE	Interpretation
Fantasy (FS)	0.941	0.659	High reliability; convergent validity achieved
Empathic Concern (EC)	0.894	0.552	Acceptable reliability and validity
Perspective Taking (PT)	0.870	0.530	Acceptable reliability and validity
Personal Distress (PD)	0.894	0.552	Acceptable reliability and validity
Average	**0.900**	**0.573**	Overall constructs demonstrate strong reliability and convergent validity

Note. All CR values exceed 0.70 and AVE values exceed 0.50, fulfilling Fornell & Larcker's (1981) criterion.

DISCUSSION

The present study aimed to adapt and validate the Interpersonal Reactivity Index (IRI) (Davis, 1983) for pre-service early childhood teachers in Indonesia. Overall, the findings confirmed that the Indonesian version preserved the original four-factor structure—Fantasy, Empathic Concern, Perspective Taking, and Personal

Distress—indicating that the multidimensional construct of empathy proposed by Davis is stable across cultural and educational contexts.

The excellent reliability coefficients (α = .976; CR = 0.87–0.94) and adequate convergent validity (AVE = 0.53–0.66) demonstrate that the adapted scale captures empathy consistently among teacher candidates. The factor structure confirmed through both EFA and CFA supports the internal coherence and theoretical distinction of the four subdimensions. The fit indices (CFI = .940; TLI = .919; RMSEA = .086) were within acceptable ranges , suggesting that the hypothesized model is empirically sound and theoretically meaningful(Abidin et al., 2020; Baharum et al., 2023).

These results also provide evidence of cross-cultural equivalence. Following (Beaton et al., 2000), semantic and conceptual equivalence were maintained throughout the adaptation process (Aiken's V = 0.80–1.00). The factorial invariance between the Indonesian and original English structures implies that the empathy construct is universal, yet adaptable to local socio-educational nuances.

From a contextual perspective, empathy plays a critical role in early childhood education. The high loading on the Empathic Concern factor aligns with the collectivist orientation of Indonesian culture, where caring and emotional attunement are embedded in teacher identity. Conversely, moderate scores on Personal Distress may indicate that pre-service teachers are still developing emotional regulation skills needed to manage intense classroom situations—an aspect central to teacher professional growth.

This adaptation also offers practical implications. The validated Indonesian IRI can be used not only to assess empathy levels but also to evaluate the impact of social–emotional training programs for pre-service teachers. Given the increasing emphasis on socio-emotional learning (Cavioni et al., 2024), the instrument provides a culturally relevant and psychometrically sound foundation for future research and professional development in early childhood education.

Finally, the strong psychometric results indicate that the IRI-Indonesian version can serve as both a diagnostic and evaluative tool in teacher education. It supports evidence-based interventions aiming to enhance empathy, improve classroom climate, and foster children's emotional well-being key components of quality early childhood teaching practice.

CONCLUSION

The present study successfully adapted and validated the Interpersonal Reactivity Index (IRI) for Indonesian pre-service early childhood teachers. The adaptation process followed Beaton et al.'s (2000) guidelines, ensuring semantic and conceptual equivalence, while the psychometric evaluation confirmed a stable four-factor model consistent with Davis's (1983) theoretical structure. The Indonesian version demonstrated excellent internal consistency (α = .976), strong reliability (CR = 0.87–0.94), and satisfactory construct validity (AVE = 0.53–0.66).

These findings indicate that empathy—as measured by the IRI—is a culturally relevant and theoretically robust construct for teacher education in Indonesia. The adapted instrument can serve as a valid tool for assessing empathy and evaluating social–emotional learning interventions in early childhood teacher preparation programs, contributing to the development of emotionally responsive and culturally grounded educators.

NOTE/NOTES

The author declares no conflicts of interest regarding this study. The research was conducted in compliance with ethical principles for human subjects, and participation was voluntary and anonymous. The adapted instrument was used with permission for academic and research purposes. Data supporting the findings of this study are available from the author upon reasonable request.

FUNDINGS

This study was conducted as part of the author's doctoral research and received no external financial support.

ACKNOWLEDGEMENTS

The author would like to express sincere appreciation to Riza Amalia for her valuable guidance in understanding and applying the stages of instrument adaptation, and to Suci Ramdaeni for kindly providing access to her class for the pilot project. The author also extends gratitude to the expert reviewers for their thoughtful evaluations of semantic and conceptual equivalence, and to all pre-service early childhood teachers who participated voluntarily in this study.

REFERENCES

1. Abidin, Z., Hotifah, Y., Suhariadi, F., & Fajrianthi. (2020). Adaptation and Validation of Entrepreneurial Intention Scale in the Santri Context in Indonesia. 714–719. https://doi.org/10.2991/assehr.k.201214.325
2. Atkins, D., Uskul, A. K., & Cooper, N. R. (2016). Culture shapes empathic responses to physical and social pain. Emotion, 16(5), 587–601. https://doi.org/10.1037/emo0000162
3. Baharum, H., Ismail, A., Awang, Z., McKenna, L., Ibrahim, R., Mohamed, Z., & Hassan, N. H. (2023). Validating an Instrument for Measuring Newly Graduated Nurses' Adaptation. International Journal of Environmental Research and Public Health, 20(4), 2860. https://doi.org/10.3390/ijerph20042860

4. Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. Spine, 25(24), 3186.
5. Cavioni, V., Broli, L., & Grazzani, I. (2024). Bridging the SEL CASEL Framework with European educational policies and assessment approaches. International Journal of Emotional Education, 16(1), 6–25. Scopus. https://doi.org/10.56300/ULTX1565
6. Cliff, N. (1988). The eigenvalues-greater-than-one rule and the reliability of components. Psychological Bulletin, 103(2), 276–279. https://doi.org/10.1037/0033-2909.103.2.276
7. 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
8. Denham, S. A., Bassett, H. H., Way, E., Kalb, S., Warren-Khot, H., & Zinsser, K. (2014). "How Would You Feel? What Would You Do?" Development and Underpinnings of Preschoolers' Social Information Processing. Journal of Research in Childhood Education, 28(2), 182–202. https://doi.org/10.1080/02568543.2014.883558
9. Hambleton, R. K., & de Jong, J. H. A. L. (2003). Advances in translating and adapting educational and psychological tests. Language Testing, 20(2), 127–134. https://doi.org/10.1191/0265532203lt247
10. Hoffman, M. (2000). Empathy and Moral Development. https://www.jstage.jst.go.jp/article/arepj1962/35/0/35_157/_article/-char/ja/
11. Hotifah, Y., Nawangsari, N. A. F., & Yoenanto, N. H. (2022). ADAPTATION OF TEACHER'S RELATIONAL COMPETENCE SCALE IN THE CONTEXT OF TEACHERS OF GUIDANCE AND COUNSELING. Jurnal Psikologi, 21(2), 177–186. https://doi.org/10.14710/jp.21.2.177-186
12. Indreswari, H., Yafie, E., & Ramalingam, K. (2022). The Effectiveness of Parental Self-Efficacy Program to Improve Children's Moral Development with Single-Parent Status. Pertanika Journal of Social Sciences and Humanities, 30(1), 381–405. https://doi.org/10.47836/pjssh.30.1.18
13. Jiang, N., Li, H., Ju, S.-Y., Kong, L.-K., & Li, J. (2025). Pre-service teachers' empathy and attitudes toward inclusive education—The chain mediating role of teaching motivation and inclusive education efficacy. PLOS ONE, 20(4), e0321066. https://doi.org/10.1371/journal.pone.0321066
14. Ledesma, R. D., Ferrando, P. J., Trógolo, M. A., Poó, F. M., Tosi, J. D., & Castro, C. (2021). Exploratory factor analysis in transportation research: Current practices and recommendations. Transportation Research Part F: Traffic Psychology and Behaviour, 78, 340–352. https://doi.org/10.1016/j.trf.2021.02.021
15. Pedersen, P. (2009). Inclusive cultural empathy: A relationship-centred alternative to individualism. South African Journal of Psychology, 39(2), 143–156. https://doi.org/10.10520/EJC98539
16. Setyawan, C. F., Sudirman, D. F., Sari, D. P., Nurulita, F. R., & Eva, N. (2021). Asesmen Perkembangan Sosio Emosinal pada Anak Usia Dini. Seminar Nasional Psikologi Dan Ilmu Humaniora (SENAPIH), 1(1), 58–70.
17. Supriatna, E., Hanurawan, F., Eva, N., Rahmawati, H., & Yusuf, H. (2024). Analyzing Factors Affecting Social Skills Development Among Students in Indonesian Schools. Islamic Guidance and Counseling Journal, 7(1). https://doi.org/10.25217/0020247447100
18. Tabachnick, B. (2007). Experimental Designs Using ANOVA.