ORIGINAL ARTICLE

Validation of the Basic Empathy Scale in a Sample of Young Adults in Kosovo: A Cross-Sectional Study

Njomza Llullaku[1] · G. H. P. van der Helm[2,3] · Jesse Roest[2]

Abstract

This cross-sectional study examines the psychometric properties of the Albanian version of the Basic Empathy Scale in Kosovo in a sample of 476 students (61.6% female; 38.4% male) aged 18 to 25 years ($M = 20.25$; $SD = 1.61$). A confirmatory factor analysis was conducted on the 20 items in R with Lavaan. The best-fitting two-factor model for the data contained 11 items for affective empathy and nine items for cognitive empathy; furthermore, it had the same factor structure as the original model. Internal consistency reliabilities were considered acceptable for cognitive ($\alpha = .68$) and affective ($\alpha = .73$) empathy. The scales correlated positively and significantly with each other ($r = .45$, $p < .01$) and gave good Cronbach's alpha reliability coefficients. The concurrent validity of the Basic Empathy Scale was confirmed, with cognitive and affective empathy scales significantly and inversely correlated with the Inventory of Callous Unemotional Traits.

Keywords BES · ICU · Validation · Reliability · Kosovo

As Kosovo is a newly formed country, data and research related to all aspects of its society are somewhat limited. The transition as a post-war country (Malcolm, 1998) reflected upheaval in several Kosovan social development issues — especially related to traditional family structures. These were undermined by numerous factors including rural to urban migration; the promotion of individualism through urban living; lack of localized institutions; lack of education, (mental) health, and law; and high unemployment levels. This context has posed a particular risk of increasing criminal behaviors among unemployed youths (Heiniger & Renaud, 2008).

✉ Njomza Llullaku
 njomza.llullaku@uni-pr.edu

[1] Department of Social Work/Faculty of Philosophy, University of Prishtina "Hasan Prishtina", Str. George Bush, No. 31, Prishtina 10000, Kosovo

[2] Leiden University of Applied Sciences, Zernikedreef 11, 2300 AJ Leiden, The Netherlands

[3] Research Institute Child Development and Education, University of Amsterdam, P.O. Box 15780, 1001 NG Amsterdam, The Netherlands

The Juvenile Justice System in Kosovo, which is relatively new, was established with the assistance of the European Union and the United Nations in 2007. Prior to this year, Kosovan justice did not prescribe any specific guidelines for the treatment of children, adolescents, and young adults; issues related to all these social groups were processed through the adult system (Hamilton & Anderson, 2008). Currently, the Assembly of Kosovo enforces the system of law through the Juvenile Justice Code, which places young adults under the age of 21 (depending on the seriousness of the crime) under the responsibility of the juvenile system (Juvenile Justice Code, 2018). In the 10 years since its establishment, Kosovo has managed and maintained a standard of care that has been acknowledged as being one of the best practices by both the EU and the United Nations (Team, 2018).

Within this context, this study aims to provide a valid means for assessing at-risk youth, helping those already placed within the system, and helping those undergoing rehabilitation in Kosovo. Specifically, through the analysis of the psychometric properties of the Basic Empathy Scale (BES) in relation to young adults in Kosovo, the BES can provide local institutions with a valid tool for empathy evaluation, and, at an international level, extend the current literature through the inclusion of Kosovo.

Empathy

A broad body of research has acknowledged empathy as a key component for molding well-adjusted and productive members in society (Bezerra et al., 2021; de Medeiros et al., 2021; Hudson et al., 2019; Llorca-Mestre et al., 2017; Samper et al., 2021). Conversely, a deficiency in empathy can result in selfishness, anti-social and aggressive behavior, and an increased likelihood of committing offences among young offenders (Bacchini et al., 2018; Bezerra et al., 2021; Coetzee, 2019; Heynen et al., 2017, 2018; Llorca-Mestre et al., 2017; Narvey et al., 2020; Van Zonneveld et al., 2017).

As such, empathy development can play an import role in the identification of at-risk youth regarding criminal offences and assist in the rehabilitation of young offenders (de Medeiros et al., 2021; Heynen et al., 2016; Jolliffe & Farrington, 2004, 2006; Samper et al., 2021) Therefore, it is imperative to have a reliable tool for empathy measurement.

Cohen and Strayer defined empathy as "the understanding and sharing in another's emotional state or context" (p. 523, 1996); they rationalized that this definition could allow for the application of two aspects of empathy: effective and cognitive. Effective empathy involves the mirroring of an emotional response that one has witnessed in another, whereas cognitive empathy involves the ability to comprehend an emotional response but not feel it.

Empathy evolves as individuals develop; therefore, it has been used for accurately predicting the future behaviors of individuals from a young age (Fricket al., 2014; Llorca-Mestre et al., 2017; Muñoz & Frick, 2012; van Zonneveld et al., 2017). Moreover, youth aged between 15 and 24 years (Fricket al., 2014; Llorca-Mestre et al., 2017; Muñoz & Frick, 2012; van Zonneveld et al., 2017) who have high

empathy levels have better relationships with others (e.g., peers, authority figures, and parents) (Heynen et al., 2016; Van Lissa et al., 2015), unlike those with low empathy, such youth build long-lasting relationships. Youth with high empathy are less likely to be impulsive or engage in risk taking behaviors (Allemand et al., 2015; Rodriguezet al., 2019). They are also often engaged in a greater number of prosocial behaviors (e.g., volunteering, helping, and sharing) (de Medeiros et al., 2021; Eisenberg & Miller, 1987; Hoffman, 2008; Hudson et al., 2019) and show greater social competencies in adulthood (Allemand et al., 2015).

Deficits in empathy, particularly during youth, are reportedly a greater factor in offending among youth compared to adults, and young adults with reduced empathy tend to commit crimes more frequently (Llorca-Mestre et al., 2017; Narvey et al., 2020; Samper et al., 2021; Wang et al., 2017). Further, those with low empathy tend to reject authority figures, and this places them at risk of deviant peer relationships (Hudson et al., 2019).

In their longitudinal study, Allemand et al. (2015) found that both positive and negative empathy-related changes among youth affected their adult coping skills 25 years later. They also found that youth who reported a decrease in empathy were more likely to have fewer friends and suffer from loneliness and negative emotionality.

In conclusion, empathy plays a crucial role in the prosocial development of young people and can play a vital role in treatment and prevention of criminal behavior in vulnerable youth (Frick & Kemp, 2021; Heynen et al., 2018; Rodriguez et al., 2019; Samper et al., 2021; van Zonneveld et al., 2017). As such, a valid and reliable measure is required for assessing empathy in youth, and this measure should be able to predict behavior risk assessments and assistance in rehabilitation.

Rationale for the Development of the Basic Empathy Scale

In the pursuit of a valid and reliable self-report questionnaire for the measurement of empathy, Jolliffe and Farrington (2004, 2006) developed the BES. It was in response to the weaknesses that they had observed and reported in the previously used measures. Specifically, these previous measures had been suitable for measuring sympathy—an issue derived from the interchangeability between the earlier literatures on sympathy and cognitive empathy, respectively. It was also found that these tools often failed to measure both cognitive and affective empathy (Jolliffe & Farrington, 2004, 2006)—two aspects that have been generally considered as being essential for defining empathy (Cohen & Strayer, 1996; D'Ambrosio et al., 2009; Heynen et al., 2016; Jolliffe & Farrington, 2004, 2006; You, Lee, & Lee, 2018).

The identification of the two aspects of empathy is crucial, as multiple studies have shown that cognitive empathy is present in individuals and children who are prone to callous and unemotional traits—that is, traits that are common in individuals that are potentially susceptible to perpetrating violent antisocial behaviors, which can often continue into adulthood (Frick & Kemp, 2021; Frick et al., 2014; Giroux & Guay, 2021; Heynen et al., 2016; Waller et al., 2020; Zych et al., 2019). This group has been shown to be deficient in affective empathy, that is, the ability to

directly share the exact emotions of others. This deficiency means that while these individuals can comprehend that their actions may cause serious distress (which can be perceived through cognitive empathy) to others, they remain emotionally indifferent to the plight of others.

The psychometric properties of the BES have been found to be adequate—to varying degrees—in different cultural contexts including Europe, America, and Asia. BES scores were positively linked with prosocial behaviors and negatively linked with certain aspects of psychopathy; consequently, it cannot be assumed that they are supported within a Kosovo context (Geng et al., 2012; Heynen et al., 2016; McLaren et al., 2019; Pechorro et al., 2015; Sanchez-Perez et al., 2014; You et al., 2018).

Methods

Participants and Design

This cross-sectional study was conducted with 476 students (38.4% male) aged 18 to 25 years ($M = 20.25$; $SD = 1.61$) from the University of Prishtina "Hasan Prishtina". The sampled students were from the following academic fields: economics (19%), law (25%), engineering (18%), psychology (18%), and sociology (18%), representing all 3 years of bachelor studies: 1st (29%), 2nd (33%), and 3rd (31%). In the case of Law studies, the program was four years, and therefore, the 4th year of study was included (7% of total participants). In terms of location, 59% of the students hailed from urban areas (Table 1).

Data Collection

Data collection took place in 2020 at the University of Prishtina "Hasan Prishtina". Being a state-run University, it offers tuition-free education along with student housing that places students from all regions of Kosovo in the capital of Prishtina. The university permits access to a range of young adults. It also provides 3 years of student housing that is characterized by mixed gender and socioeconomically diversity; furthermore, it had the necessary mechanisms for conducting ethical oversight. The student populations of the chosen departments were identified as representing a broad range of student types.

The survey, once approved by the ethics committee of Faculty of Philosophy/University of Prishtina "Hasan Prishtina", was presented to the department heads for an overview and then to teaching professors. Upon provision of consent, the students received an outline of the study in person, and surveys were distributed to the participants.

All participants voluntarily agreed to participate, signed an informed consent form, and were told that their answers would be treated confidentially and anonymously, with responses being accessible only to the researchers.

Table 1 Baseline characteristics of study participants ($N=476$)

	Number	%
Age 20.25† 1.61 ±		
Gender		
Female	293	61.6
Male	183	38.4
Field of study		
Economics	90	18.9
Law	120	25.2
Engineering	88	18.5
Psychology	89	18.7
Sociology	89	18.7
Year of study*		
1st	139	29.2
2nd	158	33.2
3rd	149	31.3
4th	30	6.3
Location*		
Urban	279	59.1
Rural	193	40.9

†*Mean ± standard deviation*

Measures

Basic Empathy Scale

The BES (Jolliffe & Farrington, 2006) is a 20-item self-report measure that assesses two components of empathy: cognitive empathy (9 items) and affective empathy (11 items); response categories are rated on a five-point Likert-type scale (ranging from $1 = I$ *strongly disagree* to $5 = I$ *strongly agree*) created in English. Upon receiving permission from the author of the BES, the English version of the BES was translated into Albanian language through multistage back-translation, a requirement for conveying the same meaning as the original text. The initial forward translation was conducted by two independent experts and was followed by reconciliation of the translation by a third independent expert, with the result being back-translated to English by a fourth independent expert who did not have access to the original text. During this comparison, no meaningful changes were found. Table 2 presents the descriptive statistics of all the items. After reverse scoring several items of the cognitive (3 items) and affective (5 items) empathy scales, higher scores were considered to be indicative of higher levels of cognitive or affective empathy. The original scale has been shown to be a valid and reliable measure for assessing cognitive and affective empathy with good internal consistency (cognitive: $\alpha = 0.79$; affective: $\alpha = 0.85$) (Jolliffe & Farrington, 2006).

Table 2 Item descriptives of the Basic Empathy Scale (*N* = 476)-Albanian language

Scale/Item (English)	Scale/Item (Albanian)	N	M	SD	Skewness	Kurtosis
Cognitive empathy	**Empatia kognitive**					
3. I can understand my friend's happiness when she/he does well at something	3. Unë e kuptoj gëzimin e shokut/shoqes tim/e kur e kryen një punë me sukses	466	4.24	0.88	−1.73	3.85
6. I find it hard to know when my friends are frightened	6. Unë e kam të vështirë ta di se kur miqtë e mi kanë frikë	467	2.47	1.04	0.68	−0.13
9. When someone is feeling "down" I can usually understand how they feel	9. Kur dikush ndihet keq apo i/e mërzitur / "për tokë," zakonisht mund ta kuptoj se si ndjehet/ndihet	473	3.75	0.99	−0.89	0.63
10. I can usually work out when my friends are scared	10. Unë zakonisht mund ta kuptoj kur miqtë e mi kanë frikë apo janë të frikësuar	465	3.79	0.85	−0.95	1.16
12. I can often understand how people are feeling even before they tell me	12. Unë shpesh mund ta kuptoj se si ndjehen njerëzit edhe para se të më tregojnë	460	3.52	1.01	−0.48	−0.24
14. I can usually work out when people are cheerful	14. Unë zakonisht e kuptoj kur njerëzit janë të gëzuar	462	3.98	0.83	−1.23	2.28
16. I can usually realize quickly when a friend is angry	16. Unë zakonisht e kuptoj shpejt kur një shok/shoqe është i/e zemëruar	469	3.83	0.92	−1.03	1.28
19. I am not usually aware of my friend's feelings	19. Unë zakonisht nuk jam i vetëdijshëm për ndjenjat e shokut/shoqes tim/sime	474	2.23	1.01	0.72	−0.02
20. I have trouble figuring out when my friends are happy	20. Unë e kam të vështirë të di se kur janë të gëzuar miqtë e mi	473	2.12	1.02	1.01	0.64
Affective empathy	**Empatia Afektive**					
1. My friend's emotions do not affect me much	1. Emocionet e miqve të mi nuk ndikojnë tek unë dhe aq shumë	476	2.68	1.17	0.36	−0.81
2. After being with a friend who is sad about something, I usually feel sad	2. Pasi që rri me një shok/shoqe që është pikëlluar/mërzitur për diçka, zakonisht pikëllohem/mërzitem edhe unë	466	3.62	0.99	−0.85	0.29
4. I get frightened when I watch characters in a good scary movie	4. Unë frikësohem kur i shoh personazhet në një film të frikshëm që është i realizuar vizualisht mire	465	2.86	1.31	0.15	−1.17
5. I get caught up in other people's feelings easily	5. Unë mund të ndikohem lehtë nga ndjenjat e të tjerëve	467	3.02	1.14	−0.10	−0.83
7. I do not become sad when I see other people crying	7. Unë nuk pikëllohem kur i shoh të tjerët duke qarë	463	2.19	1.11	0.88	0.13
8. Other people's feelings do not bother me at all	8. Ndjenjat e njerëzve të tjerë nuk më shqetësojnë aspak	467	2.10	0.99	0.87	0.53

Table 2 (continued)

Scale/Item		N	M	SD	Skewness	Kurtosis
11. I often become sad when watching sad things on TV or in films	11. Unë pikëllohem shpesh kur shoh gjëra pikëlluese në tel-evizion apo në filma	460	3.39	1.15	−0.29	−0.84
13. Seeing a person who has been angered has no effect on my feelings	13. Nëse e shoh një person që e kanë zemëruar, kjo nuk ka kur-rfarë ndikimi tek ndjenjat e mija	472	2.40	1.06	0.64	−0.17
15. I tend to feel scared when I am with friends who are afraid	15. Unë kam tendencë të kem frikë kur rri me miq të cilët janë të frikësuar,	465	2.65	1.10	0.26	−0.76
17. I often get swept up in my friend's feelings	17. Unë zakonisht ndikohem nga ndjenjat e miqve të mi	465	3.19	1.08	−0.21	−0.69
18. My friend's unhappiness does not make me feel anything	18. Pa-lumturia e një shokut/shoqes nuk më shtyn të ndjej asgjë	468	2.20	1.01	0.88	0.51

Inventory of Callous Unemotional Traits

The ICU scale (Kimonis et al., 2008) is a 24-item self-report measure designed to evaluate callous and unemotional traits in youth; response categories are rated on a four-point Likert-type scale ranging from 0 (*not at all true*) to 3 (*definitely true*). The ICU has parent, teacher, and self-report versions. The present study used the self-report version, which has been shown to be a reliable and valid measure for examining CU-traits in adolescent offenders (Kimonis et al., 2008). The scale is divided into three subscales: Callousness (e.g., "the feelings of others are unimportant to me"; $\alpha = 0.70$), Unemotional (e.g., "I hide my feelings from others"; $\alpha = 0.64$), and Uncaring (e.g., "I try not to hurt others' feelings"; $\alpha = 0.73$) (Kimonis et al., 2008). The three subscales together form a higher-order callous-unemotional dimension ($\alpha = 0.77$). The present study showed good reliability for the overall ICU-factor ($\alpha = 0.75$) as well as for the subscales of callousness ($\alpha = 0.72$) and uncaring ($\alpha = 0.68$). For the unemotional subscale, the reliability was unsatisfactory ($\alpha = 0.56$).

Demographic Survey

A demographic survey was constructed to collect data on the participants' sociodemographics. This survey included the participant variables of age, gender, field of study, year of study, and residence location (rural vs. urban).

Statistical Analysis

To examine the replicability of the two-factor structure of the original BES version, a confirmatory factor analysis (CFA) was conducted using the *Lavaan* package software (Version 0.6–10) in the R environment (Rosseel, 2012). Furthermore, an alternative factor model was examined by adding a method factor containing negatively worded items. In the Korean version of the BES, You et al. (2018) found that a factor model with a method factor indicated a superior fit to the original two-factor model. To account for non-normally distributed ordinal variables, the mean and variance-adjusted weighted least squares (WLSMV) estimation procedure were used (Li, 2016).

Currently, there is no consensus on power analyses in structural equation modeling and CFA (Kline, 2016). As a rule of thumb, a ratio of cases to variables between 10:1 and 20:1 is recommended. Given the relatively large sample size ($N = 476$), CFA was considered appropriate. Additionally, investigating the psychometric properties of the BES through CFA allowed to compare results with findings of previous BES validation studies. Assumptions of normality of the data were inspected by means of computing skewness and kurtosis values. All values were within the acceptable range of skewness between -3 and $+3$ and kurtosis

within -10 and $+10$ when utilizing SEM (Brown, 2015). Additionally, SEM is recognized as a robust analytical method regarding assumptions of normality.

The goodness-of-fit of the model was evaluated by calculating several indices: the comparative fit index (CFI), Tucker-Lewis index (TLI), goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI), and root-mean-square error of approximation (RMSEA). For the adequate model fit, cut-off values of CFI ≥ 0.90, TLI ≥ 0.90, GFI ≥ 0.90, and RMSEA < 0.08 are necessary, whereas CFI ≥ 0.95, TLI ≥ 0.95, GFI ≥ 0.95, AGFI ≥ 0.90, and RMSEA < 0.05 are indicative of good model fit (Hu & Bentler, 1999; Kline, 2016).

Cronbach's alpha was calculated to examine the internal consistency reliability of the BES scales. Furthermore, the concurrent validity of the BES scales was examined by calculating Pearson's *r* correlations between the BES scales and the ICU scales. Concurrent validity is demonstrated if the cognitive empathy and affective empathy scales are significantly and inversely correlated with the Callousness, Unemotional, and Uncaring scales of the ICU. Calculations of Cronbach's alpha and correlational analyses were conducted using SPSS version 24.

Ethics Statement

All aspects of the current research met local ethical guidelines, including adherence to the legal requirements of Kosovo. The study was approved by the Ethics Committee of Faculty of Philosophy/University of Prishtina "Hasan Prishtina".

Results

A confirmatory factor analysis (CFA) of the 20 items was conducted in R with Lavaan. The two-factor model that best fitted the data contained 11 items for affective empathy and 9 items for cognitive empathy; furthermore, it shared the same factor structure with the original model. The 20-item model showed a poor fit to the data: $\chi^2(169) = 712.13$, $p < 0.001$; CFI $= 0.811$; TLI $= 0.787$; AGFI $= 0.929$; RMSEA $= 0.094$ (90% CI $= 0.087$, 0.102). The standardized factor loadings of the items ranged from 0.39 to 0.63 for positively worded items and between -0.32 to -0.69 for negatively worded items. Notably, one item, "I tend to feel scared when I am with friends who are afraid," did not load significantly on the factor of affective empathy.

Next, the model that had an additional method factor containing negatively worded items was examined. The results indicated an adequate fit to the data: $\chi^2(159) = 428.22$, $p < 0.001$; CFI $= 0.906$; TLI $= 0.888$; GFI $= 0.978$; AGFI $= 0.963$; RMSEA $= 0.068$ (90% CI $= 0.061$, 0.076). The standardized factor loadings of the items ranged from 0.39 to 0.66 for positively worded items and between -0.32 and -0.69 for negatively worded items. The item "I tend to feel scared when I am with friends who are afraid" demonstrated a very low, but significant, standardized loading (0.11). Removing this item from the model only slightly improved the model fit. Therefore, a model containing all 20 items and an additional method

factor was accepted as the final model. A post hoc power analysis was conducted using the semPower package (Moshagen & Erdfelder, 2016) in R. The results suggested that a sample size of $N = 467$ with a power larger than $> 99.99\%$ (1—beta = 1—6.463622e − 09) to reject a wrong model (with df = 159) with an amount of misspecification corresponding to RMSEA = 0.050 on alpha = 0.05.

Internal consistency reliabilities were acceptable for cognitive ($\alpha = 0.68$) and affective ($\alpha = 0.73$) empathy. The scales correlated positively and significantly with each other ($r = 0.45$, $p < 0.01$).

In order to examine measurement invariance, the factor structure of the BES was comparable between Males and females. The model fit for the male sample ($n = 183$) was poor: $\chi^2(159) = 301.05$, $p < 0.001$; CFI = 0.787; TLI = 0.746; GFI = 0.944; AGFI = 0.905; RMSEA = 0.082 (90% CI = 0.067, 0.096); power = 0.962. However, the model fit for the female sample ($n = 293$) was adequate: $\chi^2(159) = 348.89$, $p < 0.001$; CFI = 0.897; TLI = 0.877; GFI = 0.972; AGFI = 0.953; RMSEA = 0.073 (90% CI = 0.063, 0.083); power = 0.9996). Further examination of the factor loadings of the CFA models for males and females indicated that for the female subsample, item 15 ("I tend to feel scared when I am with friends who are afraid") of the Affective empathy factor was non-significant (standardized loading = 0.01). For the male subsample, item 1 ("My friends' emotions don't affect me much") and item 7 ("I don't become sad when I see other people crying") did not load significantly on the Affective empathy factor. Therefore, a measurement invariance based on gender could not be further explored.

Concurrent Validity

Next, the concurrent validity of the BES was examined. The cognitive and affective empathy scales correlated negatively and significantly with ICU scales. Table 3 presents all the relevant means, standard deviations, and correlations.

Discussion

This current research was primarily conducted to test the psychometric properties of the Albanian version of the BES on a sample of youth from Kosovo. A reliability analysis and confirmatory factor analysis were completed; these resulted in a reliable and valid 20-item bidimensional Albanian version of the BES, which shares the factor structure of the original model (Jolliffe & Farrington, 2006). This result differs from the work of other researcher, specifically those working with delinquency, who have reported lower itemed bidimensional versions of the BES (Heynen et al., 2016; McLaren et al., 2019; Pechorro et al., 2017; Van Langen et al., 2015).

Additionally, using the Inventory of Callous Unemotional Traits scales, concurrent validity was confirmed, with both cognitive and affective empathy scales correlating negatively and significantly with ICU (Frick, 2004). The use of ICU for establishing concurrent validity for the BES is well established, with similar results having been replicated in previous research, with subjects reporting a decrease in

Table 3 Descriptive statistics and correlations between study variables

Scale	α	N	M	SD	Correlations				
					(2)	(3)	(4)	(5)	(6)
Affective empathy (1)	.73	476	3.38	0.56	.45 ($p < .001$; 95% CI=.38, .52)	−.40 ($p < .001$; 95% CI=−.47, −.32)	−.28 ($p < .001$; 95% CI=−.36, −.20)	−.35 ($p < .001$; 95% CI=−.43, −.27)	−.50 ($p < .001$; 95% CI=−.56, −.43)
Cognitive empathy (2)	.68	476	3.81	0.52		−.42 ($p < .001$; 95% CI=−.49, −.34)	−.17 ($p < .001$; 95% CI=−.26, −.08)	−.32 ($p < .001$; 95% CI=−.40, −.24)	−.46 ($p < .001$; 95% CI=−.53, −.39)
Callousness (3)	.72	476	0.61	0.51			.11 ($p = .013$; 95% CI=.02, .20)	.29 ($p < .001$; 95% CI=.21, .37)	.76 ($p < .001$; 95% CI=.72, .80)
Unemotional (4)	.56	476	1.53	0.55				.25 ($p < .001$; 95% CI=.16, .33)	.54 ($p < .001$; 95% CI=.47, .60)
Uncaring (5)	.68	476	0.88	0.50					.75 ($p < .001$)
Overall ICU (6)	.75	476	0.92	0.36					

The scales Affective and Cognitive Empathy are measured on a 5-point Likert type scale. The scales Callousness, Unemotional, and Uncaring are measured on a 4-point Likert type scale

ICU traits when there is an increase in either aspect of empathy (i.e., cognitive, and affective) (Frick & Kemp, 2021; Heynen et al., 2016; Pechorro et al., 2017).

However, upon the examination measurement invariance, differences were found in model fit for the separate male and female subsamples. Regarding the female sample, item 15 ("I tend to feel scared when I am with friends who are afraid") of the Affective empathy factor was non-significant. For the male sample, two negatively phrased items, namely item 1 ("My friends' emotions don't affect me much") and item 7 ("I don't become sad when I see other people crying") did not load significantly on the Affective empathy factor, a result also reported by other researchers (Heynen et al., 2016; Pechorro et al., 2015; Salas-Wright et al., 2012; Van Langen et al., 2015).

Heynen et al. (2016) suggest that this may be the result of misunderstanding due to low education levels; however, this study's sample was composed of university students, challenging that assumption. Alternatively, what possibly links these samples is localized interpretations of masculinity. Research has shown both a link between a masculine ideology of patriarchal societies and prisons population to destructive masculine traits such as emotional suppression, expectations of dominance/aggression, and risk-taking behaviors, traits that also place downward pressure upon empathy (American Psychological Association, 2018; Gabbiadini et al., 2016; Ingram et al., 2019; Kupers, 2005; Pemberton, 2013; Ricciardelli et al., 2015).

Limitations

As this study was conducted with a sample of students from the University of Prishtina "Hasan Prishtina" which was characterized by gender and economic diversity, all the participants had a similar education level; a factor that should be considered for generalization purposes. While the final version of the Albanian BES contained 20 items, matching the original model, one item from the CFA models for males and females from the Affective empathy factor item 15, namely, "I tend to feel scared when I am with friends who are afraid," demonstrated a very low, but significant, standardized loading (0.11). The removal of this item only slightly improved the model fit, so it was retained. In addition, owing to poor model fit in the factor structure between the male and female subsample, a measurement invariance based on gender could not be further explored.

To conclude, this paper adds to the literature on Kosovo empathy evaluation capabilities by supporting the construct validity and reliability of the Albanian version of the BES, a result in line with previous research. That said, limitations in sample diversity and poor factor structure between the male and female subsamples suggest that this should be intended for research purposes only. Therefore, prior to the potential adoption of the Albanian version of the BES as (1) a diagnostic measure in the identification of at-risk youth and (2) a successful measure in juvenile delinquency programs that aim to reduce ICU traits, additional research is needed using larger and more diverse samples, including educational, criminality, and clinical samples.

Author Contribution The authors Njomza Llullaku and G.H.P van der Helm contributed to the study conception and design. Material preparation and data collection were performed by Njomza Llullaku, and analysis was conducted by Jesse Roest. The first draft of the manuscript was written by Njomza Llullaku, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Compliance with Ethical Standards This study was approved by the University of Prishtina "Hasan Prishtina"/Faculty of Philosophy ethics committee (2168), and we certify that the study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Research Involving Human Participants All research has met local ethical guidelines for the study of human participants, including adherence to the legal requirements of Kosovo. The study was approved by the University of Prishtina "Hasan Prishtina"/Faculty of Philosophy ethics committee (2168).

Research Involving Animals Not applicable.

Informed Consent All participants in this research provided informed consent.

Competing Interests The authors declare no competing interests.

References

American Psychological Association, Boys and Men Guidelines Group. (2018, August). *APA Guidelines for Psychological Practice with Boys and Men*. Retrieved October 10, 2021, from https://www.apa.org/about/policy/boys-men-practice-guidelines.pdf

Allemand, M., Steiger, A. E., & Fend, H. A. (2015). Empathy development in adolescence predicts social competencies in adulthood. *Journal of Personality, 83*(2), 229–241. https://doi.org/10.1111/jopy.12098

Bacchini, D., De Angelis, G., & Dragone, M. (2018). Empathy and utilitarian judgment in incarcerated and community control adolescents. *Criminal Justice and Behavior, 45*(7), 1093–1109. https://doi.org/10.1177/0093854818781438

Bezerra, V. A., & d. S., Camino, C. P. d. S., Galvão, L. K. d. S., & Sampaio, L. R. (2021). Predictive variables of young people's willingness to help people at risk of suicide. *Trends in Psychology*. https://doi.org/10.1007/s43076-021-00108-9

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford publications.

Coetzee, L. (2019). Victim empathy in young sex offenders in the emergent adulthood developmental phase. *Journal of Sexual Aggression, 26*(2), 251–262. https://doi.org/10.1080/13552600.2019.1618931

Cohen, D., & Strayer, J. (1996). Empathy in conduct-disordered and comparison youth. *Developmental Psychology, 32*(6), 988–998. https://doi.org/10.1037/0012-1649.32.6.988

D'Ambrosio, F., Olivier, M., Didon, D., & Besche, C. (2009). The Basic Empathy Scale: A French validation of a measure of empathy in youth. *Personality and Individual Differences, 46*(2), 160–165. https://doi.org/10.1016/j.paid.2008.09.020

de Medeiros, M. W. A., Hattori, W. T., & Yamamoto, M. E. (2021). Priming effects on prosocial behavior of children: A systematic review. *Trends in Psychology, 29*(4), 637–654. https://doi.org/10.1007/s43076-021-00076-0

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin, 101*(1), 91–119. https://doi.org/10.1037/0033-2909.101.1.91

Frick, P. J. (2004). *Inventory of callous unemotional traits*. University of New Orleans.

Frick, P. J., & Kemp, E. C. (2021). Conduct disorders and empathy development. *Annual Review of Clinical Psychology, 17*(1), 391–416. https://doi.org/10.1146/annurev-clinpsy-081219-105809

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? *A Comprehensive Review. Psychological Bulletin, 140*(1), 1–57. https://doi.org/10.1037/a0033076

Gabbiadini, R., & P., Andrighetto, L., Volpato, C., & Bushman, B. J. (2016). Acting like a tough guy: Violent-sexist video games, identification with game characters, masculine beliefs, & empathy for female violence victims. *PLoS ONE, 11*(4), e0152121–e0152121. https://doi.org/10.1371/journal.pone.0152121

Geng, Y., Xia, D., & Qin, B. (2012). The Basic Empathy Scale: A Chinese validation of a measure of empathy in adolescents. *Child Psychiatry & Human Development, 43*(4), 499–510. https://doi.org/10.1007/s10578-011-0278-6

Giroux, S., & Guay, M. C. (2021). Assessing the contribution of callous–unemotional traits and affective empathy to aggressive behaviour among teenagers hosted in a youth protection centre. *Psychology, Crime & Law*. https://doi.org/10.1080/1068316x.2021.1929977

Hamilton, C., & Anderson, K. (2008, July 5). *Justice for children: Juvenile crime and juvenile justice practice in Kosovo*. Retrieved August 15, 2021, from https://drive.google.com/file/d/1t6GK3l7kUO5w46bmmaU3YDYx3IL4pZix/view

Heiniger, J.-p., & Renaud, R. (2008, July 5). *Prevention of juvenile delinquency in Kosovo; Psychosocial Intervention Review*. Retrieved August 20, 2021, from https://tdh-europe.org/library/prevention-of-juvenile-deliquency-in-kosovo---psychosocial-intervention-review/2315

Heynen, E. J. E., van der Helm, G. H. P., Cima, M. J., Stams, G. J. J. M., & Korebrits, A. M. (2017). The feelings of others don't impress me much — Effects of living group climate on empathy in adolescent male offenders. *Psychiatry, Psychology and Law, 24*(1), 118–127. https://doi.org/10.1080/13218719.2016.1188440

Heynen, E. J. E., Van der Helm, G. H. P., Stams, G. J. J. M., & Korebrits, A. M. (2016). Measuring empathy in a german youth prison: A Validation of the German version of the Basic Empathy Scale (BES) in a sample of incarcerated juvenile offenders. *Journal of Forensic Psychology Practice, 16*(5), 336–346. https://doi.org/10.1080/15228932.2016.1219217

Heynen, E. J. E., van der Helm, G. H. P., Wissink, I. B., Stams, G. J. J. M., & Moonen, X. M. H. (2018). "I don't care about what you want!" The Relation between juvenile delinquents' responses to social problem situations and empathy in secure juvenile institutions. *Journal of Interpersonal Violence, 33*(9), 1412–1426. https://doi.org/10.1177/0886260515618212

Hoffman, M. L. (2008). Empathy and prosocial behavior. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (1st ed., pp. 440–455). The Guilford Press.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hudson, S. K. T. J., Cikara, M., & Sidanius, J. (2019). Preference for hierarchy is associated with reduced empathy and increased counter-empathy towards others, especially out-group targets. *Journal of Experimental Social Psychology, 85*, 103871. https://doi.org/10.1016/j.jesp.2019.103871

Ingram, D., & J. P., Espelage, D. L., Hatchel, T., Merrin, G. J., Valido, A., & Torgal, C. (2019). Longitudinal associations between features of toxic masculinity and bystander willingness to intervene in bullying among middle school boys. *Journal of School Psychology, 77*, 139–151. https://doi.org/10.1016/j.jsp.2019.10.007

Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior, 9*(5), 441–476. https://doi.org/10.1016/j.avb.2003.03.001

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence, 29*(4), 589–611. https://doi.org/10.1016/j.adolescence.2005.08.010

Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Munoz, L. C., Aucoin, K. J., & Morris, A. S. (2008). Assessing callous-unemotional traits in adolescent offenders: Validation of the Inventory of Callous-Unemotional Traits. *International Journal of Law and Psychiatry, 31*(3), 241–252. https://doi.org/10.1016/j.ijlp.2008.04.002

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford publications.

Kupers, T. A. (2005). Toxic masculinity as a barrier to mental health treatment in prison. *Journal of Clinical Psychology, 4*(4), 713–724. https://doi.org/10.1002/jclp.20105

Juvenile Justice Code, 06/L –006 C.F.R. (2018). Retrieved May 12, 2021, from https://md.rks-gov.net/desk/inc/media/B5AFE545-3908-4F63-98E0-0A8DD593B499.pdf

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Llorca-Mestre, A., Malonda-Vidal, E., & Samper-García, P. (2017). Prosocial reasoning and emotions in young offenders and non-offenders. *The European Journal of Psychology Applied to Legal Context, 9*(2), 65–73. https://doi.org/10.1016/j.ejpal.2017.01.001

Malcolm, N. (1998). *Kosovo: A short history*. Macmillan.

McLaren, V., Vanwoerden, S., & Sharp, C. (2019). The Basic Empathy Scale: Factor structure and validity in a sample of inpatient adolescents. *Psychological Assessment, 31*(10), 1208–1219. https://doi.org/10.1037/pas0000741

Moshagen, M., & Erdfelder, E. (2016). A new strategy for testing structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(1), 54–60. https://doi.org/10.1080/10705511.2014.950896

Muñoz, L. C., & Frick, P. J. (2012). Callous-unemotional traits and their implication for understanding and treating aggressive and violent youths. *Criminal Justice and Behavior, 39*(6), 794–813. https://doi.org/10.1177/0093854812437019

Narvey, C., Yang, J., Wolff, K. T., Baglivio, M., & Piquero, A. R. (2021). The interrelationship between empathy and adverse childhood experiences and their impact on juvenile recidivism. *Youth Violence and Juvenile Justice, 19*(1), 45–67. https://doi.org/10.1177/1541204020939647

Pechorro, P., Kahn, R. E., Gonçalves, R. A., & Ray, J. V. (2017). Psychometric properties of Basic Empathy Scale among female juvenile delinquents and school youths. *International Journal of Law and Psychiatry, 55*, 29–36. https://doi.org/10.1016/j.ijlp.2017.10.008

Pechorro, P., Ray, J. V., Salas-Wright, C. P., Maroco, J., & Gonçalves, R. A. (2015). Adaptation of the Basic Empathy Scale among a Portuguese sample of incarcerated juvenile offenders. *Psychology, Crime & Law, 21*(7), 699–714. https://doi.org/10.1080/1068316x.2015.1028546

Pemberton, S. (2013). Enforcing Gender: The constitution of sex and gender in prison regimes. *Signs: Journal of Women in Culture and Society, 39*(1), 151–175. https://doi.org/10.1086/670828

Ricciardelli, R., Maier, K., & Hannah-Moffat, K. (2015). Strategic masculinities: Vulnerabilities, risk and the production of prison masculinities. *Theoretical Criminology, 19*(4), 491–513. https://doi.org/10.1177/1362480614565849

Rodriguez, L. M., Martí-Vilar, M., Esparza Reig, J., & Mesurado, B. (2019). Empathy as a predictor of prosocial behavior and the perceived seriousness of delinquent acts: A cross-cultural comparison of Argentina and Spain. *Ethics & Behavior, 31*(2), 91–101. https://doi.org/10.1080/10508422.2019.1705159

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Samper, P., Llorca, A., Malonda, E., & Mestre, M. V. (2021). Examining the predictors of prosocial behavior in young offenders and nonoffenders. *International Journal of Behavioral Development, 45*(4), 299–309. https://doi.org/10.1177/0165025421995930

Salas-Wright, O., & R., & Vaughn, M. G. (2013). Assessing empathy in Salvadoran high-risk and gang-involved adolescents and young adults: A Spanish validation of the Basic Empathy Scale. *International Journal of Offender Therapy and Comparative Criminology, 57*(11), 1393–1416. https://doi.org/10.1177/0306624X12455170

Sanchez-Perez, N., Fuentes, L. J., Jolliffe, D., & Gonzalez-Salinas, C. (2014). Assessing children's empathy through a Spanish adaptation of the Basic Empathy Scale: Parent's and child's report forms. *Frontiers in Psychology, 5*, 1438. https://doi.org/10.3389/fpsyg.2014.01438

Team, P. (2018, July 5). *EU and UNICEF mark 10th Anniversary of support to Justice for Children Reforms in Kosovo*. Retrieved June 17, 2021, from https://eeas.europa.eu/delegations/kosovo/48027/eu-and-unicef-mark-10th-anniversary-support-justice-children-reforms-kosovo_en

van Langen, M. A., Wissink, I. B., van Vugt, E. S., Van der Stouwe, T., & Stams, G. J. J. (2014). The relation between empathy and offending: A meta-analysis. *Aggression and Violent Behavior, 19*(2), 179–189. https://doi.org/10.1016/j.avb.2014.02.003

Van Lissa, C. J., Hawk, S. T., Branje, S. J., Koot, H. M., Van Lier, P. A., & Meeus, W. H. (2015). Divergence between adolescent and parental perceptions of conflict in relationship to adolescent empathy development. *Journal of Youth and Adolescence, 44*(1), 48–61. https://doi.org/10.1007/s10964-014-0152-5

van Zonneveld, L., Platje, E., de Sonneville, L., van Goozen, S., & Swaab, H. (2017). Affective empathy, cognitive empathy and social attention in children at high risk of criminal behaviour. *Journal of Child Psychology and Psychiatry, 58*(8), 913–921. https://doi.org/10.1111/jcpp.12724

Waller, R., Wagner, N. J., Barstead, M. G., Subar, A., Petersen, J. L., Hyde, J. S., & Hyde, L. W. (2020). A meta-analysis of the associations between callous-unemotional traits and empathy, prosociality, and guilt. *Clinical Psychology Review, 75*, 101809. https://doi.org/10.1016/j.cpr.2019.101809

Wang, X., Lei, L., Yang, J., Gao, L., & Zhao, F. (2017). Moral disengagement as mediator and moderator of the relation between empathy and aggression among Chinese male juvenile delinquents. *Child Psychiatry & Human Development, 48*(2), 316–326. https://doi.org/10.1007/s10578-016-0643-6

You, S., Lee, J., & Lee, Y. (2018). Validation of Basic Empathy Scale: Exploring a Korean version. *Current Psychology, 37*(4), 726–730. https://doi.org/10.1007/s12144-016-9554-8

Zych, I., Farrington, D. P., Nasaescu, E., Jolliffe, D., & Twardowska-Staszek, E. (2020). Psychometric properties of the Basic Empathy Scale in Polish children and adolescents. *Current Psychology*. https://doi.org/10.1007/s12144-020-00670-y

Zych, I., Ttofi, M. M., & Farrington, D. P. (2019). Empathy and callous-unemotional traits in different bullying roles: A systematic review and meta-analysis. *Trauma, Violence, & Abuse, 20*(1), 3–21. https://doi.org/10.1177/1524838016683456