

Development and empirical evaluation of a new empathy questionnaire for early adolescents

UNA MIKAC, VESNA BUŠKO, MONIKA IVANOVIĆ and ANA BABIĆ ČIKEŠ

The paper presents the construction and initial validation data on a new instrument intended to measure empathy in early adolescents. Research on adults mostly supports the hypothesis of a cognitive and an affective component of empathy, but the structure of empathy in children is less clear. In designing this instrument, empathy was conceptualized as rather stable disposition composed of the suggested cognitive and affective dimensions. The questionnaire was administered on two samples of elementary school students ($n_1 = 202$; $n_2 = 133$). Average age in both samples was 12 years, with boys and girls equally represented. The analyses, including EFA, CFA, and validity evidence based on relationships with demographic data and some personality traits, suggest a general empathy factor and a reverse coding method factor. The results are interpreted in light of existing theories and previous empirical findings published on empathy measures, with a special emphasis on methodological issues encountered.

Key words: empathy, dimensionality, early adolescents, CFA

Researchers from multiple disciplines such as counseling, child, and social psychology are interested in the concept of empathy from a practical and a theoretical viewpoint (Cotton, 1992; Duan & Hill, 1996; Miklikowska, Duriez, & Soenens, 2011; Van Lange, 2008). There are many definitions and different hypotheses concerning the structure of empathy, but most agree that it encompasses: (a) an emotional reaction similar to what another person is feeling, and (b) an understanding of other person's feelings.

Mehrabian and Epstein (1972), for example, viewed empathy as the ability to share other people's emotion and developed a measure according to this definition (Questionnaire Measure of Emotional Empathy; QMEE). Today there seems to be a consensus on empathy as a multidimensional concept consisting of both affective and cognitive components (Duan & Hill, 1996). Still, the exact nature of these components differs between theorists. Davis (1980) proposed four dimensions: two affective dimensions

that differentiate between orientation of the reaction to others (empathic concern) and to self (personal distress), one cognitive dimension (perspective taking), and a dimension concerning the ability to transpose oneself into fictional situations (fantasy). The Interpersonal Reactivity Index (IRI), developed to measure these four dimensions, is one of the frequently used measures of empathy today (Albiero, Mataricardi, Speltri, & Toso, 2009). Further research revealed a general empathy factor behind these dimensions which coincides with empathic concern, and partly with perspective taking and fantasy (Cliffordson, 2001, 2002). Baron-Cohen and Wheelwright (2004) argued that empathy is a unitary construct where affective and cognitive components are indiscernible and developed an instrument to measure this construct, the Empathy Quotient. However, subsequent research using this instrument pointed to a complex factor structure including cognitive empathy, emotional reactivity, and social skills (Dimitrijević, Hanak, Vukosavljević-Gvozden, & Opačić, 2012).

The emotional-cognitive dichotomy of empathy is supported by neuroanatomic research (Singer, 2006). The affective component, i.e., sharing another's emotions, is accompanied by the activation of the limbic system, while the cognitive components, i.e., perspective taking, are accompanied by the activation of the prefrontal and temporal cortex. These neuroanatomic findings also help clearing up the development of empathy. Processes based in the limbic system develop earlier than prefrontal cortex, i.e., emotion-related aspects of empathy appear earlier in development

Una Mikac, Department of Psychology, Faculty of Humanities and Social Sciences, Ivana Lučića 3, 10000 Zagreb, Croatia.

E-mail: umikac@ffzg.hr (the address for correspondence);

Vesna Buško, Department of Psychology, Faculty of Humanities and Social Sciences Zagreb, Croatia;

Monika Ivanović, Ministry of Science, Education and Sports of the Republic of Croatia;

Ana Babić Čikeš, Department of Psychology, Faculty of Humanities and Social Sciences, University J. J. Strossmayer, Osijek, Croatia.

than cognitive components. Already in newborns we witness crying in reaction to crying of other babies of their age, but not in reaction to a cry from an older baby, their own cry, or a synthetic cry. This may be considered a basic form of empathy, i.e., sharing other's emotional state but without cognitive understanding of the other's state (Eisenberg, Fabes, & Spinrad, 2006). After age 5 their empathy abilities increase with an increase in capacity to take perspective. There seems to be a further increase in empathy in adolescence compared to childhood age, although findings concerning changes during adolescence are less clear (Eisenberg et al., 2006). There are some indications that affective empathy does not change after the age of six, except a small drop during adolescence after which it reverts to previous levels, while cognitive and, subsequently, general empathy increase (Dadds et al., 2008; Van der Graaff et al., 2014). This is in line with the fact that some areas of prefrontal cortex involved in empathic responses mature as late as 25 years of age (Singer, 2006). After late adolescence, empathy seems to be relatively stable (Grühn, Rebusal, Diehl, Lumley, & Labouvie-Vief, 2008).

However, how empathy develops during adolescence remains unclear (Wentzel, Filisetti, & Looney, 2007). To approach this issue, a valid measure of empathy in this age is needed. Most often used methods for adults are questionnaires. When exploring empathy during early and late adolescence multiple authors also choose an age appropriate questionnaire, but frequently encounter psychometric issues. There were some attempts to adapt IRI to children, but the resulting measure did not show a stable structure (Garton & Gringart, 2005; Litvack-Miller & McDougall, 1997; Špelić & Zuliani, 2011) and had variable reliability ($.44 \leq \alpha \leq .71$). A frequently used measure is The Empathy Index for Children and Adolescents (IECA; Bryant, 1982), developed on the basis of QMEE. Its original purpose was to measure emotional aspects of empathy, but subsequent studies found a complex factor structure, with one of the factors interpreted as a cognitive factor (De Wied et al., 2007; Lasa Aristu, Holgado Tello, Carrasco Ortiz, & del Barrio Gándara, 2008). Some authors have doubts about the reliability of this factor ($\alpha \approx .60$) and the fact that all its items were reverse coded (De Wied et al., 2007). Other authors showed that, despite these issues, there was evidence for its discriminant validity compared to the affective component (Dadds et al., 2008).

Croatian version of the empathy scale for adolescents was based on IRI, QMEE, and empathy subscale of Junior Eysenck Personality Questionnaire (Eysenck, Easting, & Pearson, 1984; as cited in Raboteg-Šarić, 1993). It was composed of two factors, affective and fantasy, the former being the basis of Emotional Empathy Scale used in later research (Raboteg-Šarić, 1993). It contained some items that could be interpreted in terms of cognitive empathy, but which also referred to an emotional reaction. Therefore, there was a need to develop a measure in Croatian which

would encompass both affective and cognitive components and enable researchers to further explore the structure and development of empathy.

The purpose of this paper is to present a part of validation procedures based on two studies using a new instrument administered on early adolescents in Croatia (Buško, Babić, & Ivanović, 2008). The authors aimed to design a measure of empathy in early adolescence encompassing affective and cognitive components. First we present the results of the EFA and CFA and the data on the stability of structure across samples, gender, and age. Second, we present the reliability data and validity evidence based on relationships with variables that showed a consistent relation to empathy: gender, aggression, and emotional intelligence measures. We expected girls to score higher on all aspects of empathy (Eisenberg & Lennon, 1983; Gilet, Mella, Studer, Grühn, & Labouvie-Vief, 2013). Empathizing with a victim should inhibit aggression; therefore individuals more prone to empathizing in general should be less prone to aggressive behavior, which is confirmed by a body of studies (Björkqvist, Österman, & Kaukiainen, 2000; Miller & Eisenberg, 1988). Regarding the relation of empathy to emotional intelligence, we expected a positive correlation, based on empirical evidence and Salovey and Mayer (1990)'s statements that these constructs both include emotion appraisal and expression and contribute to the quality of interpersonal relationships (Mayer, Caruso, & Salovey, 1999; Schutte et al., 2001). We also present preliminary data on age differences in empathy components.

MATERIALS AND METHODS

Samples and procedure

The data were collected on two independent samples of 5th to 8th grade elementary school students from one big and one medium sized Croatian town ($n_1 = 202$; $n_2 = 133$; respectively). The students completed the empathy questionnaire and a set of other measures that differed between two studies (Table 5) during a single class period (45 min). The age was similar in both samples ($M_1 = 12.79$, range = 11-15; $M_2 = 12.80$, range = 10-15), with girls and boys equally represented (girls: $n_1 = 104$; $n_2 = 72$).

Measures

Empathy self-report questionnaire. The instrument was developed by the second and third author of this paper, with preliminary version composed of newly designed and items adapted from previous measures (Bryant, 1982; De Wied et al., 2007; Raboteg-Šarić, 1993). Item content was chosen as to represent both the cognitive and affective component and adapted in form to early adolescent age. The questionnaire

contained 22 items, 11 intended to measure understanding the feelings of others, e.g., "It is hard for me to understand why someone became upset", and the rest of items intended to measure sharing those feelings, e.g., "I feel sorry for downtrodden animals". The answers were given on a five point Likert-type scale ranging from 0 (*totally untrue*) to 4 (*totally true*). The total score was expressed so that higher scores represent higher empathy, with eight of the items reverse coded before that.

Aggression Scale. Aggression was assessed via the adapted Direct & Indirect Aggression Scale ($k = 9$; Björkqvist, Lagerspetz, & Österman, 1992; Milanović, 2004) which uses average peer-rating (of all the classmates) as a measure of direct, indirect, and total aggression (e.g., "Hits a pupil", "Gossips about the pupil he/she is angry with"). The ratings were given on scale from 0 (*never acts that way*) to 4 (*often acts that way*).

Emotional intelligence measures. Two composite performance-based instruments and peer-ratings using nominations method were administered. Peer-ratings of specific hypothesized emotional intelligence skills and abilities used were: the ability of the individual to recognize emotions, to cheer others up, to express emotions verbally, the capacity to show concern for others, and conflict resolutions skills (Buško & Babić, 2006). The pupils' task was to name three of his/her classmates who were the most skillful, and three who were the least skillful in each of these abilities. The total scores were calculated as $\{1 + [(number\ of\ the\ times\ chosen\ as\ the\ most\ skillful) - number\ of\ the\ times\ chosen\ as\ the\ least\ skillful] / number\ of\ the\ times\ chosen\ in\ any\ category\}$ (Papić & Kulenović, 2003), thus ranging was from zero to two with higher scores reflecting higher levels of emotional intelligence.

The adapted version of Emotional Analysis Test (Kulenović, Balenović, & Buško, 2000) is designed to measure understanding and analysis of emotions. In each of 32 items a participant was given a mixed mental state or a complex emotion and had to decide which of five combinations of emotions offered correspond to the given state, e.g., Disappointment is a mixture of (a) sadness and surprise, (b) guilt and shame, (c) sadness and shame, (d) guilt and despite, or (e) joy and despite. Total score was calculated as a sum of correct answers with higher scores indicating higher ability. Internal consistency of the scale in this study estimated by Cronbach α was .74.

The Emotion Management Test (Babić Čikeš & Buško, 2015; Buško & Babić Čikeš, 2013) is intended to measure the ability to regulate own and others' emotions in early adolescents by offering descriptions of 16 hypothetical situations that provoke different emotions. The task was to rate the usefulness of each of the four declared behavioral reactions to the situation specified. Response options were selected so to reflect possible reactions of varied efficiency in managing own or others' emotions. For example, an

item includes the situation: "Iva is very angry because her younger sister demolished her book.", and participants are asked: "How useful is each of the following behaviors for Iva to feel better: (a) to go into another room and take a deep breath, (b) to yell at her sister, (c) to tell her parents what happened, and (d) to think about doing something similar with her sisters toys?". Answers were scored from 0 to 2 points regarding their usefulness for the person in the situation, and the points summed up to form a total score with higher scores representing higher ability. Internal consistency of the scale estimated by Cronbach α in this study was $\alpha = .63$.

RESULTS

Data analyses included factor analyses of the empathy questionnaire items, both exploratory, to offer initial structure of the questionnaire and its stability across samples, and confirmatory, to test the structural hypotheses based on theoretical expectations and the outcomes of EFA. The characteristics of the final version of the questionnaire were than further analyzed including reliability indices and relationships with relevant variables. The analyses were done using SPSS v.21 and LISREL 8.80 (Jöreskog & Sörbom, 2006).

Exploratory factor analysis

EFA was performed on each of the samples. Answers on the items did not show univariate or multivariate normality as indicated by visual inspection and statistical tests ($\chi^2_{skewness\ and\ kurtosis} = 988.450, p < .001$). Because asymmetry and kurtosis were of similar direction, Pearson correlations and principal axis factoring were deemed as acceptable representation of the item relationships. The parallel analysis (O'Connor, 2000) indicated that first two factors explained more variance than expected by chance on both samples (95% CI), while the third factor was statistically significant only in Sample 2. The first two factors in the two- and the three-factor solution (principal axis factoring with oblimin rotation) in both samples consisted of the same items, while the third factor varied in its composition and showed a correlation with the first ($r_{Sample\ 1} = -.314; r_{Sample\ 2} = .307$) and not the second factor ($r_{Sample\ 1} = -.102; r_{Sample\ 2} = -.089$). Therefore, we decided for a two-factor solution, which explained about 27% of the total variance in each sample (Table 1). One of the two factors could be classified as having high congruence between samples (Tucker's congruence coefficient, $\Phi_{F1} = .965$) and the other as fairly similar on two samples ($\Phi_{F2} = .892$) according to Lorenzo-Seva and ten Berge (2006). The two factors were independent ($r_{F1, F2}$: -.005 and .030 for Samples 1 and 2, respectively) and both included the items originally intended to measure affective and cognitive component. The structure was almost the same when the analysis was performed separately

Table 1
Structure of the empathy questionnaire resulting from principal axis factoring with oblimin rotation

Item	Sample 1 (n = 202)		Sample 2 (n = 133)	
	F1	F2	F1	F2
E19. I'm happiest when the people around me are happy. ^a	.71		.62	
E1. When I see that one of my friends is sad, I also get sad. ^a	.66		.63	
E16. When I see someone wanting to cheat or deceive someone else, I wish to protect him. ^b	.59		.63	
E14. Sometimes I imagine how the children without parents feel. ^b	.59		.65	
E17. I am sorry for the downtrodden animals on the road. ^a	.59		.55	
E7. I rarely notice that there is someone in the park who does not have anybody to play with. ^{a, c}	.58		.57	
E21. I get angry when someone offends or beats someone weaker than him/herself. ^a	.56		.59	
E9. It upsets me when I see an animal harmed. ^a	.54		.59	
E13. I hate it when someone treads flowers or plucks plants. ^a	.51		.47	
E10. It bugs me when the teacher is unfair to someone from the class. ^b	.48		.51	
E5. I feel terrible if I have to tell somebody some bad news. ^a	.47		.50	
E6. When someone gets me angry, I try to imagine why he did it. ^b	.46		.46	
E2. Before I criticize others, I consider how I would feel in that situation. ^b	.45		.61	
E15. I cry when I'm saddened by song lyrics. ^a	.44			
E4. When someone gets down, it's hard for me to understand why. ^{b, d}		.61		.44
E11. It's funny to me is when someone is very shy when they find themselves in a new company. ^{a, d}		.44		.42
E18. I do not understand how someone can cry with happiness. ^{b, d}		.41		.53
E3. It's strange to me when I see that someone has stage fright before the show or during oral exam in school. ^{a, d}		.32		.46
E22. It is silly is to treat animals with respect. ^{b, d}		.31		.62
E8. I enjoy watching someone opening a gift and being happy about it. ^{b, d}		.32		
E12. Sometimes I say something rash, which can hurt other people's feelings. ^{b, c}			.34	
E20. It's hard for me to imagine how the characters from the books that I read feel. ^{b, d}				.54
Rotation sums of squared loadings	4.38	1.35	4.41	1.61

Note. Only items with a loading of .30 and higher are presented.

^a Item originally intended to measure affective empathy. ^b Item originally intended to measure cognitive empathy. ^c Excluded from further analysis.

^d Reverse coded.

for boys and girls on the aggregated data. The first factor seemed to reflect general empathy, while the second was composed entirely of reverse coded items. Due to student's feedback on item's complex content and poor psychometric characteristics two items were excluded from further analysis (Table 1).

Confirmatory factor analysis

Based on the similarity of descriptive statistics and the structure found in the EFA, CFA was performed on data from both samples aggregated together ($N = 335$) as to reach greater stability of parameter estimates. Due to departures of item distributions from normality, robust maximum likelihood estimation was used. Based on theoretical assumptions and the results of EFA we specified four models to be tested with CFA: (a) the originally conceived model including cognitive and affective factors consisting of 11 items each, with the reverse coded items also loading on a method

factor (model A in Table 2); (b) a model including a general empathy factor and a reverse coding factor (model B); (c) a model including a general empathy factor without the reversed coded factor (model C); and (d) a model including a general empathy factor with the reversed coded factor specified as independent of the general empathy factor, i.e., with reverse coded items loading only on the reversed coded factor (model D). When comparing models A and B, which are directly comparable with the only difference being the restriction of correlation of two empathy factors to unity in model B, the model fit did not change significantly, thus speaking in favor of a general empathy factor (Table 2). The differences in specification introduced in Model C and D as compared to Model B they are nested under significantly worsened the fit (Table 2).

The general empathy and reverse coding factor model, however, did not fit the data convincingly well (Table 2). Modification indices suggested adding a residual covariance parameter between two items and content analysis showed

Table 2
 Fit indices of the empathy questionnaire structure models tested with CFA

Models	χ^2	SB χ^2	df	RMSEA	RMSEA 90 % CI	CFI	$\Delta\chi^2$ (Δdf)
A. Cognitive and affective empathy with RCF	417.99	342.07	163	0.057	0.049; 0.066	0.95	
B. General empathy with RCF	417.99	341.06	164	0.057	0.048; 0.065	0.95	0 ^{a, d} (1)
C. General empathy	660.91	649.43	170	0.092	0.084; 0.099	0.85	6613.41 ^{b, c} (6)
D. General empathy and an independent RCF	432.23	354.24	170	0.057	0.049; 0.065	0.94	14.48 ^{b, d} (6)
E. General empathy and a fantasy factor with a RCF	301.35	249.27	158	0.042	0.031; 0.051	0.97	76.40 ^{b, d} (6)
F. General empathy and an independent fantasy factor with a RCF	463.67	355.35	164	0.059	0.051; 0.068	0.94	146.59 ^{c, d} (6)

Note. All χ^2 are significant at $p < .01$. RCF = reverse coding factor.

^a Compared to model A. ^b Compared to model B. ^c Compared to model E. ^d Original scaled difference test computed according to Bryant & Satorra (2012).

^e New scaled difference test computed according to Bryant & Satorra (2012) because of negative value of the original test.

they both included empathizing with animals. Therefore, we decided to test another model that would include sharing and understanding feelings of non-human and fictitious entities, similar to Fantasy scales of Davis (1980) and Raboteg-Šarić (1993). This model (model E) included general empathy factor and a fantasy factor (including items concerning empathizing with animals and books and songs content) with a reverse coding factor. This model had a significantly better fit than model B and overall satisfying fit indices (Table 2). Specifying a model where items loaded either on the general empathy factor or on the fantasy factor significantly worsened the fit (model F).

These comparisons indicated that the model with the general empathy factor and reverse coding and fantasy factor showed the best fit (model E in Table 2). However, inspection of parameter estimates showed that only two items had significant loadings on the fantasy factor and therefore, we specified it as a residual covariance, and not a separate factor. The final model, with all the insignificant parameters fixed to zero, is shown in Table 3 ($\chi^2 = 319.25$, Satorra-Bentler $\chi^2 = 261.90$, RMSEA = .041, RMSEA 90 % CI = [.031, .051], CFI = .097, AGFI = .89). Multigroup analyses were performed to test the stability of this structure across samples, gender, and age (younger, 10-12 years, $n = 140$, and older participants, 13-15 years, $n = 195$, Table 4). Analyses indicated that factors were composed of same items with equal saturations in all groups (i.e., showed metric or weak factorial invariance, Cheung & Rensvold, 2002). Two age subgroups also had equal item error variances (i.e., showed measurement error invariance), while girls and participants from the smaller town tended to have lower or equal item error variances.

Table 3
 Standardized parameter estimates for the best fitting model of structure of the empathy questionnaire

Item	Loadings		Residual variance
	General empathy	Reverse coding	
E1	.66 ^a		.57
E2	.53		.72
E3		.54 ^a	.71
E4		.61	.62
E5	.50		.75
E6	.49		.76
E8	.58		.67
E9	.49		.76
E10	.48		.77
E11		.51	.74
E13	.50		.75
E14	.62		.61
E15	.38		.86
E16	.61		.63
E17	.53		.72
E18	-.11 ^b	.50	.74
E19	.69		.53
E20	-.15	.44	.78
E21	.57		.68
E22		.49	.76

Note. Variances of both independent factors equal 0.75. Residual covariance between E9 and E17, standing for the fantasy factor, equals 0.49. All parameters significant at $p < .01$, except ^b at $p < .05$.

^a Unstandardized paths fixed to 1.

Table 4

Comparison of empathy questionnaire structure models for girls ($n = 176$) and boys ($n = 159$), younger (10-12 years, $n = 140$) and older participants (13-15 years, $n = 195$), and two samples ($n_1 = 133$, $n_2 = 202$)

	Gender							Age groups					Samples				
	df	Δdf	χ^2	SB χ^2	RMSEA	CFI	$\Delta\chi^2$	χ^2	SB χ^2	RMSEA	CFI	$\Delta\chi^2$	χ^2	SB χ^2	RMSEA	CFI	$\Delta\chi^2$
Invariance models																	
Configural invariance	334		551.35	444.43	0.045	0.96		539.35	445.64	0.045	0.97		596.4	461.38	0.048	0.96	
Weak metric invariance	355	21	569.42	457.78	0.042	0.97	15.07 ^a	575.91	470.30	0.044	0.97	30.94 ^a	636.03	485.75	0.047	0.96	31.26 ^a
Equal error variances	375	20	741.22	576.94	0.057	0.93	78.74 ^{b*}	616.30	495.32	0.044	0.96	19.24 ^b	689.08	544.52	0.052	0.94	125.15 ^{b*}

Note. All χ^2 are significant at $p < .01$. All $\Delta\chi^2$ are original scaled difference tests computed according to Bryant and Satorra (2012).

^a Compared to configural invariance model. ^b Compared to weak metric invariance.

* $p < .05$.

Both the EFA and CFA indicated one general empathy factor and one narrower factor encompassing reverse coding items and interpreted therefore as a specific method factor. The content of this method factor was clearly mixed of affective and cognitive items, and showed no correlation to general empathy factor, which would be expected of any two measures of empathy. Therefore, further analyses will be presented for the newly formed General Empathy Scale composed of items with positive saturation with general empathy factor ($k = 14$, Table 3). This scale's reliability, as indicated by Cronbach alpha, was .85 for Sample 1, and

.84 for Sample 2. The scores on the General Empathy Scale encompassed the whole theoretical range with negatively asymmetric distribution (Table 5).

Relationships with other variables

The scores on General Empathy Scale were compared across samples, gender, and four most represented age groups ($n_{11} = 57$, $n_{12} = 82$, $n_{13} = 81$, $n_{14} = 98$). Girls ($M = 43.33$, $SD = 8.89$) scored higher on general empathy than

Table 5

Descriptives for General Empathy Scale and measures of aggression and emotional intelligence and their correlations with empathy

Variable	Sample	$M (SD)$	Range	K-S (p)	Skewness (SE)	Kurtosis (SE)	Correlation with General Empathy (p)
General Empathy Scale		40.67 (9.86)	0-56	.097 (.000)	-1.1 (0.13)	1.55 (0.27)	
	A	39.11 (10.14)	0-56	.087 (.001)	-0.97 (0.17)	1.24 (0.34)	
	B	43.05 (8.94)	2-55	.131 (.000)	-1.36 (0.21)	2.67 (0.42)	
Aggression							
Direct	A	0.79 (0.56)	0.01-2.91	.089 (.000)	1.08 (0.17)	1.11 (0.34)	-.18 (.014)
Indirect	A	0.59 (0.43)	0.02-2.22	.121 (.000)	1.05 (0.17)	0.78 (0.34)	-.02 (.762)
Total	A	0.73 (0.49)	0.05-2.31	.081 (.003)	1.05 (0.17)	1.05 (0.34)	-.13 (.041)
Emotional intelligence							
Emotional Analysis Test	B	17.22 (4.93)	4-29	.076 (.055)	-0.17 (0.21)	-0.11 (0.42)	.30 (.000)
Emotion Management Test	B	64.7 (9.25)	35-84	.072 (.089)	-0.52 (0.21)	0.62 (0.42)	.42 (.000)
Peer-ratings							
Recognizing emotions	A	1.04 (0.74)	0-2	.137 (.000)	-0.04 (0.17)	-1.38 (0.34)	.30 (.000)
Concern for others	A	1.07 (0.78)	0-2	.181 (.000)	-0.12 (0.17)	-1.48 (0.34)	.33 (.000)
Cheering others up	A	0.98 (0.71)	0-2	.139 (.000)	-0.03 (0.17)	-1.31 (0.34)	.18 (.012)
Ability of verbal expression of emotions	A	1.07 (0.68)	0-2	.116 (.000)	-0.12 (0.17)	-1.19 (0.34)	.33 (.000)
Conflict resolutions skills	A	0.98 (0.74)	0-2	.166 (.000)	-0.02 (0.17)	-1.34 (0.34)	.18 (.012)

Note. A - $n = 202$; B - $n = 133$. K-S = Kolmogorov-Smirnov test.

boys ($M = 37.51$, $SD = 10.17$), $F(1, 302) = 21.887$, $p < .001$, partial $\eta^2 = .068$. The overall effect of age was significant, $F(3, 302) = 2.661$, $p = .048$, partial $\eta^2 = .026$, although none of the Scheffe's post hoc tests were significant and no linear trend was visible, with groups of 12- and 14-years-old having the lowest values ($M_{11} = 43.40$, $SD_{11} = 8.24$; $M_{12} = 39.45$, $SD_{12} = 9.89$; $M_{13} = 41.27$, $SD_{13} = 8.88$; $M_{14} = 39.21$, $SD_{14} = 11.37$). There was also a significant difference between the two study samples, $F(1, 302) = 12.652$, $p < .001$, partial $\eta^2 = .040$, with the participants from the smaller town having higher results ($M = 42.99$, $SD = 9.04$) than those from the bigger town ($M = 38.97$, $SD = 10.20$). No interaction effect was found significant. Departures from normality were found in most of the used continuous variables as shown by Kolmogorov Smirnov test and skewness and kurtosis indicators (Table 5). Because the deviations observed, although not sizable, were in different directions, Spearman's rho correlations were calculated to test the relationship with aggression and emotional intelligence. Weak but significant correlations of expected sign were found with the direct and total aggression scores. Somewhat higher, albeit still modest positive correlations were found with all measures of emotional intelligence used (r 's in range .18-.42; Table 5).

DISCUSSION

Factor analyses of the empathy questionnaire data produced a general factor of empathy and one narrower, interpreted as a method factor due to reverse coding. The observed structure was replicated on two independent study samples and on different gender and age groups.

Internal structure

The 14 items positively loading on the general factor formed the new General Empathy Scale. The computed internal consistency measure of this scale ($\alpha = .84-.85$) is among the higher in the field of measuring empathy in children (cf. De Wied et al., 2007, Garton & Gringart, 2005; Lasa Aristu et al., 2008; Litvack-Miller & McDougall, 1997; Raboteg-Šarić, 1993; Špelić & Zuliani, 2011). The affective and cognitive components could not be discerned which may have to do with the specific age of examinees. Empathy measured with various instruments shows a complex structure when examined in adults (e.g., Cliffordson, 2001) and older adolescents (Dimitrijević et al., 2012; Jolliffe & Farrington, 2006). When the structure was tested in early adolescents, some factors were interpreted as cognitive, but were either composed exclusively of reverse coded items (Dadds et al., 2008) and had questionable reliability and validity (De Wied et al., 2007; del Barrio, Aluja, & García, 2004; Lasa Aristu et al., 2008), or included affective items (Garton & Gringart, 2005; Litvack-Miller & McDougall, 1997). In early adolescence cognitive abilities underly-

ing empathy are starting to develop (Dadds et al., 2008; De Wied et al., 2007) and may still be too intertwined with the affective processes to be discerned. Further research could aim at exploring the structure of empathy from early adolescence to adulthood, and also inspect the relationship of the General Empathy Scale to other measures of affective and cognitive empathy in different stages of adolescence.

An important prerequisite when examining age differences in empathy and the development of different aspects of empathy is a sound instrument for children. In our case, the reverse coding of the items appeared as significant source of construct irrelevant variance. The general empathy factor and the reverse coding factor were mostly independent, with only two items out of possible six loading significantly on both factors. Both of these items had a low and negative loading with the empathy factor. Different authors dealt differently with the problem of the effect of reverse coding when developing empathy questionnaires in children. Some authors decided to remove the reverse coded items during development of the instrument due to lack of understanding shown by children and low item-total correlations (Garton & Gringart, 2005; Raboteg-Šarić, 1993). Retaining such items resulted in a factor consisting of reverse coded items, which was interpreted as a cognitive component due to its content (Dadds et al., 2008; De Wied et al., 2007; del Barrio et al., 2004; Lasa Aristu et al., 2008). The evidence concerning validity of those scales was limited, usually pointing to gender differentiation in the expected direction. The exception was the study of Dadds and colleagues (2008) in which cognitive scale was related as expected to children's behavioral and emotional problems, prosocial behavior, and verbal IQ. The ratings of empathy of children in study of Dadds and colleagues (2008) were made by parents and not children themselves, which lowers comparability to other findings.

It is unclear why the reverse coding had such a strong influence in our case and seemed to minimize common variance between the reverse coded items and other indicators of empathy. A possible explanation is that some participants showed an acquiescence bias, which would lead to a negative correlation with the empathy factor, while others answered in accordance with the content, which might lead to a positive relationship, with the two effects evening each other out. Yet, children might have faced problems understanding these items, although other not reverse coded items also included complex sentences. Of special interest here might be the two items that showed significant saturations with both factors.

We also tested the existence of a fantasy factor in the structure of our data. We defined this factor somewhat more broadly than Davis (1980) and Raboteg-Šarić (1993), as a factor reflecting not only the tendency to transpose oneself into fictional situations including fictional characters (e.g., books, songs), but also empathizing with non-human entities, such as animals, supposing there was an additional ele-

ment of imaging when empathy is not triggered by another real person. These items, however, did not form a separate factor. Our analysis suggests that there is a similar mechanism behind empathizing with fictional and non-fictional people, although a specific additional source of individual differences might exist when empathizing with animals. This should be explored further, with a larger number of items referring to animals, to test if this is a separate factor explaining a significant amount of variance in individual differences or is it explicable as a method factor due to similarity of content.

Analysis of structure in different groups indicated the similarity in two different samples, implying the stability of structure when measured in different settings (e.g., town size, other instruments used), and weak metric age and gender invariance. This serves as a validity evidence for the scale in terms of measurement invariance over different genders/age groups. This allows us to make inferences about the underlying dimension based on the comparison of the results of those groups on the General Empathy Scale. More unique variance, however, was measured with items when answered by boys and participants from bigger towns and their results may be more influenced by factors other than empathy and reverse coding.

Relationships with other variables

The direction and the size of correlations of selected variables with General Empathy Scale were as expected based on theoretical expectations and previous research. Girls scored higher than boys, with the effect size similar to that observed in previous research (e.g., $\eta^2 = .06$, Gilet et al., 2013). Empathy was also related to lower aggression, and the size of the correlation corresponded to the common correlation found in Miller and Eisenberg's (1988) meta-analysis for questionnaire measures ($r = -.18$). The exception was indirect aggression which was not significantly related to empathy. According to some authors understanding others' emotions is a necessary prerequisite for indirect aggression, and therefore empathy might be related to higher indirect aggression, while sharing others' emotion might lead to decrease in indirect aggression (Kaukiainen et al., 1999). It seems that understanding and sharing other emotions are indiscernible at this age, at least with our instrument, and this might lead to non-existence of the relation between empathy and indirect aggression.

Higher empathy was related to higher emotional intelligence, with most of the correlations being close to $r = .30$. This effect size is expectable for measures of two theoretically related constructs both including perception and appraisal of emotions, and still differing in methods of assessment used (self-report, peer-rating, and performance-based). Accordingly, lower correlations were obtained with Cheering others up and Conflict resolutions skills ($r = .18$), measures that differ in both of these aspects. They are peer-

ratings designed to reflect the highest level of emotional intelligence, i.e., not only perception and appraisal of emotions, but also management of one's own and others' emotion, which includes directive behavior and social skills (Mayer et al., 1999).

An unexpected result was the difference in average empathy level in two study samples. We might speculate this is due to town size, since the town with the lower mean score is the biggest in Croatia, with the population of about 800 000, while the smaller town with about 55 000 inhabitants might have a tighter community where people are more encouraged to take notice of others and their emotions. The finding should however be re-examined including data on other community features except size.

Generally, the observed relationships are similar to those found in previous research. The effect sizes follow the expected pattern, e.g., empathy is most highly correlated with measures of most similar dimension of emotional intelligence. They are of moderate but expectable size, based on theoretical interrelations of empathy to these constructs and having in mind the differences in methods used and their reliability. More data in support of its validity is needed, however, e.g., the relations to other uni- or multidimensional measures of empathy and known covariates such as prosocial behavior and social competence.

Empathy and age

This research aimed to develop a measure which would help us learn more about development of empathy, however, the preliminary results are inconclusive in this regard. Presented analyses indicate that the structure of this empathy questionnaire does not change with age with no indication of differentiation of general empathy as observed in adults (Albiero et al., 2009; Dimitrijević et al., 2012; Duan & Hill, 1996). The hypothesized changes possibly happen in older age, hence, further studies are called for to examine the developmental dynamics of empathy structure throughout the adolescence age.

Although latent structure appears stable, average empathy scores do not show an obvious trend, despite the main effect of age indicating the presence of certain variations in our data. Thus, research including wider age range, age measured in months instead of years, and better representation of different age groups might give more insight into the changes in the structure and level of empathy. Possible covariates could also be interesting, such as cognitive development, which might explain individual differences in the onset of differentiation.

Conclusions

The construction of empathy questionnaire for children resulted in a 14-item General Empathy Scale. The scale

shows reasonably high reliability indices, comparable to other sound instruments in the field. Its validity is supported by the study findings on gender differences, negative relations to aggression, and positive relations to emotional intelligence. Still, the general empathy factor might partly be due to specific operationalization used and/or the developmental stage of the study samples, this being an issue that demands further investigation.

REFERENCES

- Albiero, P., Matricardi, G., Speltri, D., & Toso, D. (2009). The assessment of empathy in adolescence: A contribution to the Italian validation of the "Basic Empathy Scale." *Journal of Adolescence*, 32(2), 393–408. doi:10.1016/j.adolescence.2008.01.001
- Babić Čikeš, A. & Buško, V. (2015). Emocionalna inteligencija u ranoj adolescenciji: korelati sposobnosti upravljanja emocijama i predikcija školskog uspjeha [Emotional intelligence in early adolescence: Correlates of emotion management ability and the prediction of school achievement]. *Društvena istraživanja*, 24(1), 21-45. doi:10.5559/di.24.1.02
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34(2), 163–175. doi:10.1023/B:JADD.0000022607.19833.00
- Björkqvist, K., Lagerspetz, K. M. J., & Kaukiainen, A. (1992). Do girls manipulate and boys fight? Developmental trends regarding direct and indirect aggression. *Aggressive Behaviour*, 18, 117-127. doi: 10.1002/1098-2337(1992)18:2<117::AID-AB2480180205>3.0.CO;2-3
- Björkqvist, K., Österman, K., & Kaukiainen, A. (2000). Social intelligence - empathy = aggression? *Aggression and Violent Behaviour*, 5, 191–200. doi:10.1016/S1359-1789(98)00029-9
- Bryant, B. K. (1982). An Index of Empathy for Children and Adolescents. *Child Development*, 53(2), 413–425. doi:10.2307/1128984
- Bryant, F. B., & Satorra, A. (2012). Principles and practice of scaled difference chi-square testing. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(3), 372-398. doi:10.1080/10705511.2012.687671
- Buško, V., & Babić, A. (2006). Prilog empirijskoj provjeri uloge emocionalne inteligencije u školskom postignuću osnovnoškolaca [An empirical evaluation of the role of emotional intelligence in explaining school achievement of primary school pupils]. *Educational Sciences*, 8(2), 313–327.
- Buško, V., & Babić Čikeš, A. (2013). Emotional intelligence in early adolescence: Validation data based on peer ratings and an objective ability-based test. *International Journal of Humanities and Social Science Invention*, 5, 54-62.
- Buško, V., Babić, A., & Ivanović, M. (2008). Prilog validaciji konstrukta emocionalne inteligencije u uzorku osnovnoškolaca [Validation of emotional intelligence construct in an elementary school sample]. In Vulić-Prtorić, A., Čubela Adorić, V., Proroković, A., Sorić, I. & Valerjev, P. (Eds.), *XVI. Dani psihologije u Zadru: Sažetci radova* (p. 45). Odjel za psihologiju Sveučilišta u Zadru.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233-255. doi:10.1207/S15328007SEM0902_5
- Cliffordson, C. (2001). Parents' judgments and students' self-judgments of empathy: The structure of empathy and agreement of judgments based on the interpersonal reactivity index (IRI). *European Journal of Psychological Assessment*, 17(1), 36. doi:10.1027//1015-5759.17.1.36
- Cliffordson, C. (2002). The hierarchical structure of empathy: Dimensional organization and relations to social functioning. *Scandinavian Journal of Psychology*, 43, 49–59. doi:10.1111/1467-9450.00268
- Cotton, K. (1992). *Developing empathy in children and youth*. Northwest Regional Educational Laboratory. Retrieved from <http://www.antelopespringscounseling.com/documents/articles/EmpathyChildrenYouth.pdf>
- Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D. J., Vassallo, S., Bunn, P., ... Masry, Y. E. (2008). A Measure of Cognitive and Affective Empathy in Children Using Parent Ratings. *Child Psychiatry and Human Development*, 39(2), 111–122. doi:10.1007/s10578-007-0075-4
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.
- De Wied, M., Maas, C., van Goozen, S., Vermande, M., Engels, R., Meeus, W., ... Goudena, P. (2007). Bryant's Empathy Index. *European Journal of Psychological Assessment*, 23(2), 99–104. doi:10.1027/1015-5759.23.2.99
- del Barrio, V., Aluja, A., & García, L. F. (2004). Bryant's Empathy Index for children and adolescents: psychometric properties in the Spanish language. *Psychological Reports*, 95(1), 257–262. doi:10.1017/sjp.2016.44
- Dimitrijević, A., Hanak, N., Vukosavljević-Gvozden, T., & Opačić, G. (2012). Psychometric properties of the Serbian version of the Empathy Quotient (S-EQ). *Psihologija*, 45(3), 257–276. doi:10.2298/PSI1203257D

- Duan, C., & Hill, C. E. (1996). The current state of empathy research. *Journal of Counseling Psychology, 43*(3), 261–274. doi:10.1037/0022-0167.43.3.261
- Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial Development. In N. Eisenberg & W. Damon (Eds.), *Handbook of child psychology. Vol 3. Social, emotional, and personality development. 6th edition* (pp. 646-718). New York: Wiley.
- Eisenberg, N., & Lennon, R. (1983). Sex Differences in Empathy and Related Capacities. *Psychological Bulletin, 94*(1), 100–131. doi:10.1037/0033-2909.94.1.100
- Garton, A. F., & Gringart, E. (2005). The Development of a Scale to Measure Empathy in 8-and 9-Year Old Children. *Australian Journal of Educational & Developmental Psychology, 5*, 17–25.
- Gilet, A.-L., Mella, N., Studer, J., Grün, D., & Labouvie-Vief, G. (2013). Assessing dispositional empathy in adults: A French validation of the Interpersonal Reactivity Index (IRI). *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement, 45*(1), 42–48. doi:10.1037/a0030425
- Grün, D., Rebucal, K., Diehl, M., Lumley, M., & Labouvie-Vief, G. (2008). Empathy across the adult lifespan: Longitudinal and experience-sampling findings. *Emotion, 8*(6), 753–765. doi:10.1037/a0014123
- Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence, 29*(4), 589–611. doi:10.1016/j.adolescence.2005.08.010
- Jöreskog, K., & Sörbom, D. (2006). *LISREL 8.80*. Scientific Software International.
- Kaukiainen, A., Björkqvist, K., Lagerspetz, K., Österman, K., Salmivalli, C., Rothberg, S., & Ahlbom, A. (1999). The Relationships Between Social Intelligence, Empathy, and Three Types of Aggression. *Aggressive Behavior, 25*, 81–89. doi:10.1002/(SICI)1098-2337(1999)25:2<81::AID-AB1>3.0.CO;2-M
- Kulenović, A., Balenović, T., & Buško, V. (2000). Test analyze emocija: Jedan pokušaj objektivnog mjerenja sposobnosti emocionalne inteligencije [Emotional analysis test: Attempt at designing an objective measure of a component of emotional intelligence]. *Suvremena Psihologija, 3*(1-2), 27–48.
- Lasa Aristu, A., Holgado Tello, F. P., Carrasco Ortiz, M. A., & del Barrio Gándara, M. V. (2008). The structure of Bryant's Empathy Index for Children: A cross-validation study. *The Spanish Journal of Psychology, 11*(2), 670–677.
- Litvack-Miller, W., & McDougall, D. (1997). The structure of empathy during middle childhood and its relationship to prosocial behavior. *Genetic, Social & General Psychology Monographs, 123*(3), 303–325.
- Lorenzo-Seva, U., & Ten Berge, J. M. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 2*(2), 57–64. doi:10.1027/1614-2241.2.2.57
- Mayer, J. D., Caruso, D., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27*, 267-298. doi:10.1016/S0160-2896(99)00016-1
- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*(4), 525–543.
- Miklikowska, M., Duriez, B., & Soenens, B. (2011). Family roots of empathy-related characteristics: The role of perceived maternal and paternal need support in adolescence. *Developmental Psychology, 47*(5), 1342–1352. doi:10.1037/a0024726
- Milanović, A. (2004). *Povezanost različitih vrsta agresivnosti i sociometrijskog statusa kod djece osnovnoškolske dobi* [The relation of different types of aggression to sociometric status of elementary school children] (Unpublished thesis). Faculty of Humanities and Social Sciences, University of Zagreb, Croatia.
- Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin, 103*(3), 324.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*(3), 396–402. doi:10.3758/BF03200807
- Papić, M., & Kulenović, A. (2003) Može li se emocionalna inteligencija prepoznati u procjenama? [Can emotional intelligence be assessed by ratings?] In D. Bratko, V. Kolesarić, & D. Maslić Seršić (Eds.), *XVI. Dani Ramira Bujasa: Sažeci priopćenja* [XVI. Ramiro Bujas' Days: Book of abstracts] (p. 57). Faculty of Humanities and Social Sciences, University of Zagreb, Croatia.
- Raboteg-Šarić, Z. (1993). *Empatija, moralno rasuđivanje i različiti oblici prosocijalnog ponašanja* [Empathy, moral reasoning and various forms of prosocial behaviour] (Unpublished thesis). Faculty of Humanities and Social Sciences, University of Zagreb, Croatia.
- Salovey, P. & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9*, 185-211.
- Schutte, N. S., Malouff, J. M., Bobik, C., Conston, T., Greeson, C., Jedlicka, C., ... Wendorf, G. (2001). Emotional intelligence and interpersonal relations. *Journal of Social Psychology, 141*, 523–536. doi:10.1080/00224540109600569
- Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews, 30*(6), 855–863. doi:10.1016/j.neubiorev.2006.06.011

- Špelić, A., & Zuliani, Đ. (2011). Uloga empatije u socijalizaciji djece s teškoćama u razrednim sredinama [Role of empathy in socialization of children with disabilities in the classroom]. *Hrvatska revija za rehabilitacijska istraživanja*, 47(2), 96–108.
- Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., & Meeus, W. (2014). Perspective taking and empathic concern in adolescence: Gender differences in developmental changes. *Developmental Psychology*, 50(3), 881–888. doi:10.1037/a0034325
- Van Lange, P. A. M. (2008). Does empathy trigger only altruistic motivation? How about selflessness or justice? *Emotion*, 8(6), 766–774. doi:10.1037/a0013967
- Wentzel, K. R., Filisetti, L., & Looney, L. (2007). Adolescent prosocial behavior: The role of self-processes and contextual cues. *Child Development*, 78(3), 895–910. doi:10.1111/j.1467-8624.2007.01039.x

