Psychometric Properties of the Basic Empathy Scale in Polish children and adolescents

The Basic Empathy Scale is widely used to measure cognitive and affective empathy in different age groups. Although empathy is studied throughout the world, research on this important psychological construct in Eastern European populations needs to be increased. In order to accomplish this, validated instruments to measure empathy are needed in this geographic area. This study was conducted to analyze the psychometric properties of the Basic Empathy Scale in Poland. The sample included 1,052 children and adolescents aged 9 to 16, enrolled in Primary and Middle schools. This was a cross-sectional study conducted with a survey that was filled in by the participants during their regular classroom hours. Confirmatory Factor Analyses was conducted, together with concurrent validity analyses and comparisons between younger and older participants and between females and males were undertaken. A final 12-item version of the Basic Empathy Scale was obtained with affective empathy and cognitive empathy factors. The Polish version of the scale showed good psychometric properties. Females scored higher on affective, cognitive and total empathy than males. Younger male participants scored higher on affective, cognitive and total empathy than older male participants. This validated measure of empathy in Polish children and adolescents can be used to study the relation between empathy and both antisocial and prosocial behaviors. In addition, this measure will allow for Poland to be included in cross-country comparisons of empathy and also used to evaluate programs focused on enhancing empathy in Poland.

Keywords: Affective empathy, Cognitive empathy, Psychometric properties, Children, Adolescents
Empathy is a personal competence that contributes greatly to the initiation and maintenance of desirable interpersonal relationships. Empathy has received great attention in psychological research and practice, and it was found to be related to desirable moral and social functioning (Romera, Casas, Gómez-Ortiz, & Ortega-Ruiz, 2019). Empathy has been defined as an emotional response related to other people’s emotions and situation that is congruent with other people’s emotional states (Eisenberg, Shea, Carlo, & Knight, 1991). Traditionally, empathy is defined by two dimensions: cognitive empathy, described as understanding other people’s emotions, and affective empathy, described as experiencing other people’s emotional states (Davis, 1983; Jolliffe & Farrington, 2006). Thus, empathy includes both thinking and feeling in relation to another’s emotional state.

Low empathy is an important predictor of antisocial behaviors, and high empathy is an important predictor of prosocial behaviors. Recent meta-analyses showed that perpetrators of bullying (Zych, Ttofi et al., 2019) and perpetrators of cyberbullying (Zych, Baldry et al., 2019) scored lower on both cognitive and affective empathy (see Zych, Farrington, & Ttofi, 2019 for a review of both meta-analyses). Similarly, other meta-analyses have found that low empathy was related to an increased likelihood of offending (Jolliffe & Farrington, 2004; van Langen, Wissink, van Vugt, Van der Stouwe, & Stams, 2014).

Regarding prosocial behaviors, meta-analytic findings confirm that children who defend the victims of bullying score high on empathy (Nickerson, Aloe, & Werth, 2015; Zych, Ttofi et al., 2019). In addition, in an experimental manipulation, it was found that empathy-induction using an emotional story was related to significantly higher cooperation with other people (Batson & Moran, 1999).
Although the causal relationship between empathy and prosocial and antisocial behaviors still needs to be confirmed, increasing empathy is frequently considered to be an important component of interventions designed to prevent or reduce antisocial behaviors such as conduct problems (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). For example, in their meta-analyses including 18 randomized controlled trials of empathy training programs, van Berkhout and Malouff (2015) found that empathy training successfully increased later levels of empathy (Hedge’s g = .51), but this was only clearly the case for select populations (university students and health professionals).

In order to examine the relation of empathy to theoretically relevant constructs in different cultures, and changes in empathy which may be the result of empathy-enhancement programs in these cultures, the development of appropriate tools to measure empathy is an essential initial step. The Basic Empathy Scale (BES), designed and validated by Jolliffe and Farrington (2006), is probably the most popular instrument to measure empathy worldwide. According to a recent review focused on the measurement of empathy, up to 2017 (Basto-Pereira & Farrington, 2019), this 2006 BES validation study has received almost 800 citations. The original scale consists of 20 items, 11 measuring affective empathy and 9 measuring cognitive empathy. The original instrument, validated with 363 English adolescents, showed good psychometric properties (Jolliffe & Farrington, 2006). Example of studies on psychometric properties of the BES in different countries are shown in table 1.

---

Insert Table 1

---
As displayed in Table 1, most of the studies show that females score higher than males on both cognitive and affective empathy. Many studies conducted with children and adolescents found an adequate fit of the original two-factor structure of the BES, including 20 items (Albeiro et al., 2009; D’Ambrosio et al., 2009; Čavojová et al., 2012; Jolliffe & Farrington, 2006; Pechorro et al., 2017). Nevertheless, most of these studies found fit indices that are commonly considered adequate but not excellent. The recommended cutoff values for indices such as CFI and TLI are close to .95, with errors (e.g., RMSEA) close to .06 (Hu & Bentler, 1999). Among the two-factor structure with 20 items, only Albiero et al. (2009) and Pechorro et al. (2017; with adolescent students but not with the incarcerated female subsample) met these strict cutoff criteria. Thus, the two-factor model with 20 items shows good psychometric properties but it seems possible to find a model with a better fit, although the model fit depends greatly on the sample used in each study.

Several studies tested different factor structures of the BES. Among them, a three-factor model with 20 items was described (Herrera-López et al., 2017). Also, two-factor models with a reduced number of items were found to have a good fit (Geng et al., 2012; Heynen, et al., 2016; Pechorro et al., 2017, Sánchez-Pérez, Fuentes, Jolliffe, & González-Salinas, 2014). Among these studies, the strict cutoff criteria for an excellent fit (Hu & Bentler, 1999) were met for a 12-item version with a two-factor structure that showed the best fit (Heynen et al., 2016) and for a 16-item version with a two-factor structure (Geng et al., 2012). Geng et al. (2012) used a sample aged 9 to 18 years old, an age range similar to the participants of the current study. Thus, it is possible that a shorter version of the BES could be more appropriate in certain settings.

The original version of the BES (Jolliffe & Farrington, 2006) includes 12 positively worded items and eight reverse worded items. The short version developed by
Heynen et al. (2016) with only two reversed items, using a sample of prisoners, showed
the best fit indices. Importantly, the authors of this study suggested that some of the
reversed items might have not been understood by the participants. Although combining
positive and reversed items is common in psychological research, new developments in
psychometrics suggest that this might negatively affect the psychometric properties of
questionnaires. Suárez-Alvarez et al. (2018) conducted a repeated measure study, in
which they administered a self-efficacy questionnaire with positive only, reversed only,
and combined-item forms to general population adults. In comparison to the positive only
and reversed only forms, the combined-item form with reversed and positive items
showed lower discrimination indices, and also had lower reliability and worse fit indices
in a confirmatory factor analysis. There seemed to be no differences regarding
acquiescence response bias between the three forms. As a result, the authors
recommended not using positive and reversed items combined in the same questionnaire.

It is possible that the use of only positively-worded items would improve the
psychometric properties of the BES. Moreover, it is important to validate the BES in the
geographic areas where research on empathy is at an earlier stage of development. To our
knowledge, the BES has not been validated or used in Poland. At the same time, the rates
of bullying in Poland are frequently found to be higher than in other countries
(Twardowska-Staszek, Zych, & Ortega-Ruiz, 2018; Zych, Ortega-Ruiz, & Llorent, 2017).

There are only a few measures of empathy that have been validated in Poland.
Among them, Kliszcz, Nowicka-Sauer, Trzeciak, Nowak, and Sadowska (2006) validated
the Jefferson Scale of Empathy (Hojat et al., 2002). The Jefferson Scale of Empathy is
specifically designed to measure empathy of health care providers with their patients.
Jankowiak-Siuda et al. (2017) have recently validated the Polish version of the Empathy
Quotient (Wakabayashi et al., 2006). This study was conducted with participants aged 15
to 80, and the psychometric properties of the scale for children and adolescents were not specifically tested. Some studies in Poland used an Internet application which induced either affective or cognitive empathy, finding that cognitive empathy was related to defending the victims of cyberbullying (Barlińska, Szuster, & Winiewski, 2018). These studies did not include self-report measures of empathy.

The BES is one of the most frequently used measures of empathy in the world, but it has not been validated in Poland. Empathy is an important psychological construct (e.g., Cohen & Strayer, 1996), and it is essential to have a measure validated for Polish children and adolescents. This would make it possible to obtain a measure that can be used to evaluate programs to enhance empathy, and to understand if the relation of empathy with prosocial and antisocial behaviors in Poland is similar to the relations among these variables found in other countries. Therefore, the objective of this study is to analyze the psychometric properties of the Polish version of the BES using a broad sample of children and adolescents. The psychometric properties of the questionnaire will be tested using the 20-item version and a short version that includes only the positively-worded items. Given that empathy is related to desirable moral functioning (e.g., Romera et al., 2019), social and emotional competencies and prosocial behaviors (e.g., Nickerson et al., 2015) and low involvement in bullying (e.g., Zych, Ttofi, et al., 2019) concurrent validity is tested through correlations with these constructs. It was hypothesized that the Polish version of the BES would have good psychometric properties.

**Method**

**Participants**

A total number of 1,052 students aged 9 to 16 years (M = 12.53, SD = 1.98; 54.4% females) participated in this study. Schools were located in the Lesser Poland geographic area, four of these were in a large city and two were in smaller towns. Students were
enrolled in four Primary Schools, Grades 4 to 7 (N = 580) and two Middle Schools, Grades 2 and 3 (N = 472). All the participants were Caucasian with Polish nationality.

**Procedure**

This was a cross-sectional study conducted using a survey method-approach. The BES was translated into Polish by the first author (a Polish native speaker), reviewed by the last author (also a Polish native speaker) and back-translated into English by an official translation service. The final versions were compared, and minor disagreements resolved. The minor disagreements were related to vocabulary that could mostly be interpreted as synonyms.

Schools were selected by convenience through the head teachers who were invited to participate in this study. These schools were then contacted, and all agreed to participate in the survey. Within each classroom, students were informed about the objective of this study by a researcher and asked to fill in the survey. Participation was voluntary and anonymous, and participants had the right to decline or withdraw from the study at any point. Students filled in a pen and paper survey during their regular classroom hours, supervised by the researchers of this project who delivered and collected the questionnaires. None of the students declined to participate or withdrew their consent. The study met all national and international ethical standards, including the Declaration of Helsinki and the data protection regulations.

**Instruments**

Empathy was measured using the Basic Empathy Scale (Jolliffe & Farrington, 2006) translated into Polish. The original and the Polish version of the BES use a 5-point Likert response scale ranging from 1 (totally disagree) to 5 (totally agree). The original scale contains 20 items: 11 items focused on affective empathy and 9 items focused on cognitive empathy. The final Polish version included 12 positive-worded items, with six
focused on affective empathy (e.g., feeling sad after being with a friend who was sad) and six focused on cognitive empathy (e.g., understanding friend’s happiness). The instrument showed good psychometric properties described in the results section.

Social and emotional competencies were measured using the Social and Emotional Competencies Questionnaire (SEC-Q) by Zych, Ortega-Ruiz, Muñoz-Morales, and Llorent (2018). This instrument (α = .90, in the current sample) contains 16 items with a 5-point Likert response scale ranging from 1 (strongly disagree) to 5 (strongly agree). It includes four factors: Self-awareness (α = .78; e.g., “I am aware of the thoughts that influence my emotions”), Self-motivation and management (α = .77; e.g., “I pursue my objectives despite the difficulties”), Social-awareness and prosocial behavior (α = .79; e.g., “I usually know how to help others who need that”) and Responsible decision making (α = .78; e.g., “I usually consider advantages and disadvantages of each option before I make decisions”). The CFA showed a good fit of the current data to this four-factor model (S/B $\chi^2 = 291.1784; df = 98; p < .001; \text{NFI} = .98; \text{NNFI} = .98; \text{CFI} = .99; \text{RMSEA} = .047; 90\% \text{CI} = .041 - .053$).

Moral disengagement was measured using The Mechanisms of Moral Disengagement Scale (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996). This instrument (α = .93) includes 32 items with a 5-point Likert response scale ranging from 1 (totally disagree) to 5 (totally agree) distributed across 4 domains: Dehumanization (α = .78; 8 items, e.g., “Some people deserve to be treated as animals”), Minimizing consequences (α = .67; 4 items, e.g., “Teasing someone does not really hurt them”), Reconstruction (α = .79;12 items, e.g., “It is alright to beat someone who bad mouths your family”), and Disconnecting agency (α = .76; 8 items, e.g., “If kids are living under bad conditions they cannot be blamed for behaving aggressively”). The CFA confirmed the four factor structures showing an adequate fit of the current data (S/B $\chi^2 = 1879.7569;$
\[ df = 458; \ p < .001; \ NFI = .90; \ NNFI = .92; \ CFI = .92; \ RMSEA = .091; \ 90\% \ CI = .087 - .096). \]

A short 19-item version was developed for Primary Education after eliminating the items that were difficult to understand for Primary School children, with a 5-point Likert response scale ranging from 1 (totally disagree) to 5 (totally agree) distributed in three domains: Dehumanization (\( \alpha = .82; \ 7 \) items), Minimizing consequences (\( \alpha = .74; \ 4 \) items), and Reconstruction (\( \alpha = .83; \ 8 \) items). The CFA for the three-factor model indicates a good fit of the data (S/B \( \chi^2 = 664.4819; \ df = 149; \ p < .001; \ NFI = .97; \ NNFI = .97; \ CFI = .98; \ RMSEA = .085; \ 90\% \ CI = .079 - .092). \]

Data analysis

First, a Confirmatory Factor Analysis (CFA) with maximum likelihood, robust method and polychoric correlations (Satorra-Bentler chi-square) was performed with EQS 6.2. This was done with a 20-item model based on the original version of the BES (Jolliffe & Farrington, 2006). Model fit was tested taking into account a combination of different indices such as the Normed Fit Index (NFI) (\( \geq .90 \)), the Non-Normed Fit Index (NNFI) (\( \geq .90 \)), the Comparative Fit Index (CFI) (\( \geq .90 \)) and the Root Mean Square Error Approximation (RMSEA) (\( \leq .08 \)) (Bentler, 1990). Factor loadings were examined and items with low factor loadings (lower than .20) were eliminated, and another Confirmatory Factor Analysis was run. Item-total correlation matrix was examined to check if all the items of the scale were correlated in the expected direction. Items with non-significant correlations and correlations in an unexpected direction were eliminated obtaining a short version of the BES with positively-worded items only. Another Confirmatory Factor Analysis was run and fit indices of all the models were compared to choose the model with the best fit.
Descriptive statistics were calculated using the PASW Statistics 20 software. Cronbach’s alphas were calculated for each factor and the total scale. Pearson item-total correlations and interitem correlations were also calculated. Concurrent validity was tested using Pearson correlations among the BES, Moral disengagement and Social and Emotional Competencies. Empathy was expected to be related to low Moral disengagement and high level of Social and Emotional Competencies. Differences between males and females, and younger versus older participants, were tested using the Student's t-test. To estimate the construct reliability, composite reliability (CR) and average variance extracted (AVE) were computed. The cut-off points used for these indices are usually .70 for CR and .50 for AVE.

Results

Factor structure and items of the BES in Polish children and adolescents

A confirmatory factor analysis using the original 20-item two-factor structure of the BES showed a poor fit to the data for Primary Education participants (S/B chi-square = 794.52, df = 169, p < .01, CFI = .69, NFI = .64. RMSEA = .09) and for Middle Education participants (S/B chi-square = 1528.12, df = 169, p < .01, CFI = .66, NFI = .63. RMSEA = .14). Thus, other models were tested to find a better fit to the data.

Factor loadings of items 1, 6 and 13 (all negatively-worded) were below .20 in both Primary and Middle Education subsamples. Model fit improved after eliminating these items for Primary Education participants (S/B chi-square = 777.68, df = 118, p < .01, CFI = .88, NFI = .86. RMSEA = .11) and for Middle Education participants (S/B chi-square = 972.77, df = 118, p < .01, CFI = .76, NFI = .74, RMSEA = .13). Nevertheless, the model fit was still poor.
Item-total correlation matrix and interitem correlation analyses showed some further concerns regarding the negatively-worded items. Item 7 (negative) had nonsignificant correlations with items 9 and 10; item 8 (negative) had nonsignificant correlations with items 4, 9 and 11; item 19 (negative) had nonsignificant correlations with item 17, and item 20 (negative) had nonsignificant correlations with items 4, 5, 11, 15 and 17. In the Middle Education sample, all the negative worded items had loadings below .40. In the Primary Education sample, negatively-worded items such as 7, 8, and 20 had loadings below .40. Thus, a model without the negatively-worded items was tested.

Alternatively, 12-item model without the negatively-worded items showed a good fit to data in Primary Education (S/B chi-square = 177.14, df = 53, p < .01, CFI = .97, NFI = .96, RMSEA = .07) and in Middle Education (S/B chi-square = 203.02, df = 53, p < .01, CFI = .94, NFI = .93, RMSEA = .08). In this case, all the factor loadings were above .40 (see figure 1). Thus, based on the fit indices and theoretical basis, the two-factor model with 12 items was considered the best.

Insert figure 1

Reliability of the BES in Polish children and adolescents

The Polish version of the BES had very good Cronbach´s alpha values for affective (Primary α = .75, Middle α = .76), cognitive (Primary α = .84, Middle α = .77) and total empathy (Primary α = .85, Middle α = .84). The Average Variance Extracted in the Primary School sample was .47 and in the Secondary School sample was .41. The Composite Reliability in the Primary School sample was .91 and in the Middle School sample was .89.
Table 2 shows that affective, cognitive and total empathy were higher in Primary Education compared to Middle Education. It also shows that affective, cognitive and total empathy were higher for females compared to males. Gender differences were consistent in both Primary and Secondary Education. Nevertheless, the lower levels of empathy in Primary compared to Secondary education was only significant for males. Thus, older males have less empathy, but there were no differences for females.

Insert table 2

Relations among empathy, moral disengagement and social and emotional competencies

Concurrent validity was tested by checking if empathy was related to social and emotional competencies and to moral disengagement. These relations were statistically significant and were in the expected direction, showing that empathy was positively related to high social and emotional competencies and low moral disengagement (see table 3). The strongest relations were found between high affective empathy and high social awareness and prosociality, and between high cognitive empathy and a high total score in social and emotional competencies.

Insert table 3

Discussion

Empathy is an important social and emotional skill that is related to low antisocial behavior (Jolliffe & Farrington, 2004) and high prosocial behavior (Nickerson et al.,
As a result, many social and emotional learning programs target increasing empathy as a way of increasing desirable behavior and social cohesion (Durlak et al., 2011). Nevertheless, research focused on empathy is not equally advanced around the world and there are some geographic areas, such as Poland, where studies on empathy are urgently needed.

The objective of this study was to test the psychometric properties of the BES in Poland using a broad sample of Polish children and adolescents. The BES is one of the most popular measures of empathy in the world (see Basto-Pereira & Farrington, 2019 for a review). This measure had not been validated in Poland. To our knowledge, this is the first study that validated a measure of empathy in Polish children and adolescents. Thus, we believe that the current results are useful in advancing knowledge about empathy in an understudied geographic area.

In this study, an original 20-item version of the BES (Jolliffe & Farrington, 2006) was tested, but this factor structure did not show a good fit to the data. Based on other studies that used a shorter version (e.g., Geng et al., 2012; Heynen, et al., 2016), and on statistical analyses of the items, a 12-item version of the BES was produced. Some studies suggest that the inclusion of both positively-worded items and negatively-worded items in a questionnaire decreases its psychometric properties (Suárez Álvarez et al., 2018). The final Polish version of the BES, with the best psychometric properties, includes only positively-worded items. Thus, it is possible that the participants had difficulties in understanding the negatively-worded items or the Likert response-scale for these items. The Polish 12-item version of the BES showed very good psychometric properties for Primary School children and Middle School adolescents.

Some gender differences were found regarding affective and cognitive empathy in Polish children and adolescents. Females scored higher than males in Primary and
Middle school subsamples, in affective and cognitive empathy. Perceived affective and cognitive empathy were stable in females but decreased with age in males. Empathy was found to be related to theoretically similar constructs such as low moral disengagement and high social and emotional competencies. Nevertheless, these relations were stronger in males than in females. Previous research found that females tend to have more advanced perceptions of social and emotional competencies (Zych et al., 2018), and therefore, it is possible that they can distinguish between empathy and similar constructs whereas males treat them all as a single, less complex construct. Future studies should examine these possibilities.

Given that the rates of antisocial behaviors such as bullying in Poland are relatively high (Twardowska-Staszek et al., 2018), programs to decrease these behaviors are urgently needed. Increasing empathy should be a component of these programs and the current study is especially useful for the evaluation of these programs, that should be specifically adapted to Polish culture. For example, compared to other countries, Poland is a country with medium-high individualistic culture and medium-high emotional expressivity endorsement (Matsumoto et al., 2008). Thus, it is reasonable to suggest that programs to promote empathy in Poland should be based on both individual and social values, and can use relatively high, but not exaggerated, expressions of empathy based on emotional expressivity culture in Poland. Future studies could also include cross-national comparisons to study similarities and differences in the relation between empathy and antisocial or prosocial behaviors in Poland and other geographic areas. This could be useful for the prevention and intervention in these behaviors.

This study has several strengths, such as the use of a broadly validated questionnaire with a large sample of Polish children and adolescents. Nevertheless, it also has some limitations. Given that the Polish version of the BES only includes positively-
worded items, it could be useful to conduct future studies that control for social desirability or acquiescence response bias. Future research could also use empathy measures that do not use self-reports, for example, other-reports focused on behavioral expressions of empathy. Despite these limitations, the current study is an important step towards filling the gaps in knowledge regarding empathy in Poland.

Given that research on empathy is not equally advanced around the world, and studies in Poland were urgently needed, the current study is an important contribution to the field. The Polish version of the Basic Empathy Scale has good psychometric properties and it can become a very useful tool for researchers and practitioners in psychology in Poland.

**Conflict of interest**

All the authors declare that they have no conflict of interest.

**Compliance with Ethical Standards**

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.
References


<table>
<thead>
<tr>
<th>Study and country</th>
<th>Participants</th>
<th>Number of items and factors</th>
<th>Psychometric properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albeiro et al. (2009)</td>
<td>655 adolescents aged 15.13 years (SD = .89), 252 males and 403 females.</td>
<td>20 items with a 2-factor structure as in Jolliffe and Farrington (2006).</td>
<td>Confirmatory Factor Analysis: NFI = .93, CFI = .95; RMSEA = .068; SRMR = .05; GFI = .91; AGFI = .88. Cronbach’s α = .87 (affective α = .86, cognitive α = .74) Females scored higher than males (stronger in affective empathy).</td>
</tr>
<tr>
<td>D’Ambrosio et al. (2009)</td>
<td>446 adolescents aged 14.8 years (SD = 1.14), 250 females and 196 males. A retest was conducted with 153 adolescents.</td>
<td>20 items with a 2-factor structure as in Jolliffe and Farrington (2006).</td>
<td>Confirmatory Factor Analysis: χ²(169) = 482.14, p &lt; 0.01, CFI = 0.78, SRMSR = 0.063, RMSEA = 0.068, 90% CI = [0.061-0.075], GFI = 0.90. Females scored higher than males in affective, cognitive and total empathy. Good test-retest reliability 3-weeks later (total r = 0.66, affective r = 0.70, cognitive r = 0.54).</td>
</tr>
<tr>
<td>Čavojová, Sirotová, &amp; Belovičová (2012)</td>
<td>426 students (215 males, 210 females) aged 10-16 years.</td>
<td>20 items with a 2-factor structure as in Jolliffe and Farrington (2006).</td>
<td>Confirmatory Factor Analysis: χ² (169) = 518.82, p &lt; .001; SRMSR = 0.072; RMSEA = 0.070; GFI = 0.884; AGFI = 0.856. Females scored higher than males in affective, cognitive and total empathy.</td>
</tr>
<tr>
<td>Geng, Xia, &amp; Qin (2012)</td>
<td>1,524 children and adolescents aged 9–18 years (48.6% females).</td>
<td>Initially: 20 items with a 2-factor structure as in Jolliffe and Farrington (2006). Final version: 16 items with a 2-factor structure.</td>
<td>Confirmatory Factor Analysis with 20 items: χ² = 1,049.31, df = 169, χ²/df = 6.21, RMSEA = .063, NFI = .79, NNFI = .79, CFI = .81, GFI = .93, AGFI = .91. Confirmatory Factor Analysis after eliminating 4 items: χ² = 186.03, df = 89, χ²/df = 2.09, RMSEA = .038, NFI = .91, NNFI = .93, CFI = .95, GFI = .97, AGFI = .95. Females scored higher than males in affective, cognitive and total empathy. Older participants scored higher than younger participants. Good test-retest correlations for cognitive (r = .60) and affective (r = .70) subscales. Good Cronbach’s alphas (cognitive α = .72, affective α = .73 total α = .77)</td>
</tr>
<tr>
<td>Study</td>
<td>Sample Description</td>
<td>Methodology</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Herrera-López, Gómez-Ortiz, Ortega-Ruiz, Jolliffe, &amp; Romera (2017)</td>
<td>747 Spanish adolescents aged 12 to 17 years (51.3% females).</td>
<td>Initially: 20 items with a 2-factor structure as in Jolliffe and Farrington (2006). Final version: A three factor structure. Confirmatory Factor Analysis with the original 2-factor structure: $S-B=839.355$; $\chi^2_S-B/(169)=4.967$; $p&lt;0.001$; NNFI=0.904; CFI=0.915; RMSEA=0.069 (90%CI 0.063 to 0.075); SRMR=0.088; AIC=968.962.</td>
<td></td>
</tr>
<tr>
<td>Heynen, Van der Helm, Stams, &amp; Korebrits (2016) Germany</td>
<td>94 incarcerated young males, 14-26 years old.</td>
<td>Initially: 20 items with a 2-factor structure as in Jolliffe and Farrington (2006). Final version: 12 items with a 2-factor structure 20 items with a 2-factor structure. Confirmatory Factor Analysis with 20 items: $\chi^2=618.163$; $X^2_S-B/(167)=3.702$; $p&lt;0.001$; NNFI=0.919; CFI=0.929; RMSEA=0.063 (90% CI 0.058 to 0.069); SRMR=0.077; AIC=284.163.</td>
<td></td>
</tr>
<tr>
<td>Jolliffe &amp; Farrington (2006) The UK</td>
<td>363 adolescents (194 males, 169 females), Mean age = 14.8 (SD = 0.48)</td>
<td>Initially: 20 items with a 2-factor structure as in Jolliffe and Farrington (2006). Final version: 12 items with a 2-factor structure 20 items with a 2-factor structure. Confirmatory Factor Analysis with 20 items: GFI = 0.89, AGFI = 0.86, RMS = 0.06. Females scored higher than males in affective and cognitive empathy. Good concurrent validity.</td>
<td></td>
</tr>
<tr>
<td>Pechorro, Kahn, Gonçalves &amp; Ray (2017) Portugal</td>
<td>377 females aged 14–19 years among which 103 were incarcerated and 274 were a school sample.</td>
<td>Initially: 20 items with a 2-factor structure as in Jolliffe and Farrington (2006). Final: 20 items in the school sample, 18 items in incarcerated females. Confirmatory Factor Analysis with 20 items in the school sample: $S-B \chi^2/df = 2.83$, IFI = 0.96, CFI = 0.96 RMSEA (90%CI) = 0.08(0.07–0.09) AIC = 138.18. Confirmatory Factor Analysis with 20 items in incarcerated females: $S-B \chi^2/df = 2.40$, IFI = 0.84, CFI = 0.84, RMSEA (90%CI) = 0.12(0.10–0.13), AIC = 65.93. Confirmatory Factor Analysis with 18 items in incarcerated females: $S-B \chi^2/df = 2.53$, IFI = 0.87, CFI = 0.86, RMSEA (90%CI) = 0.12(0.11–0.14), AIC = 68.96.</td>
<td></td>
</tr>
<tr>
<td>Sánchez-Pérez et al. (2014)</td>
<td>290 children (145 males, 145 females) aged 8–12 years</td>
<td>Initially: 20 items with a 2-factor structure as in Jolliffe and Farrington (2006). Final: 18 items with a 2-factor structure</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Good Cronbach’s alphas (school: affective $\alpha = 0.85$, cognitive $\alpha = 0.93$, total $\alpha = 0.90$; incarcerated affective $\alpha = 0.81$, cognitive $\alpha = 0.86$, total $\alpha = 0.82$). Confirmatory Factor Analysis with 18 items: $\chi^2/df = 1.52$, CFI = 0.93, TLI = 0.92, RMSEA = 0.04, SRMR = 0.07. Acceptable Cronbach’s alpha (affective $\alpha = 0.66$, cognitive $\alpha = 0.70$). No gender differences were found.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Gender and Primary versus Secondary Education differences in Affective and Cognitive Empathy

<table>
<thead>
<tr>
<th></th>
<th>Primary Education</th>
<th>Secondary Education</th>
<th>Primary vs. Secondary (total)</th>
<th>Primary vs. Secondary (males)</th>
<th>Primary vs. Secondary (females)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>Males</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Affective</td>
<td>17.28 (5.11)</td>
<td>20.52 (4.97)</td>
<td>19.12 (5.26)</td>
<td>7.22**</td>
<td>16.07 (5.19)</td>
</tr>
<tr>
<td>Cognitive</td>
<td>22.25 (5.41)</td>
<td>23.23 (4.77)</td>
<td>22.83 (5.06)</td>
<td>2.17*</td>
<td>20.35 (5.04)</td>
</tr>
<tr>
<td>Total</td>
<td>39.51 (9.35)</td>
<td>43.73 (8.72)</td>
<td>41.91 (9.19)</td>
<td>5.14**</td>
<td>36.65 (8.85)</td>
</tr>
</tbody>
</table>

* p < .05, ** p < .01; minimum score in affective and cognitive empathy = 6, maximum score = 30; minimum score in total empathy = 12, maximum score = 60
Table 3. Pearson correlations between Empathy (final scales), Moral Disengagement and Social and Emotional Competencies

<table>
<thead>
<tr>
<th></th>
<th>Affective Empathy</th>
<th>Cognitive Empathy</th>
<th>Total Empathy</th>
<th>Moral Disengagement</th>
<th>Total Social Emotional Competencies</th>
<th>Self-awareness and motivation</th>
<th>Responsible decision making</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affective Empathy</strong></td>
<td>1</td>
<td>.54**</td>
<td>.89**</td>
<td>-.13**</td>
<td>.04</td>
<td>.00</td>
<td>-.02</td>
</tr>
<tr>
<td><strong>Cognitive Empathy</strong></td>
<td>.54**</td>
<td>1</td>
<td>.86**</td>
<td>-.16**</td>
<td>.31**</td>
<td>.22**</td>
<td>.15**</td>
</tr>
<tr>
<td><strong>Total Empathy</strong></td>
<td>.87**</td>
<td>.88**</td>
<td>1</td>
<td>-.16**</td>
<td>.20**</td>
<td>.14**</td>
<td>.08</td>
</tr>
<tr>
<td><strong>Moral Disengagement</strong></td>
<td>-.12*</td>
<td>-.15**</td>
<td>-.16**</td>
<td>1</td>
<td>-.33**</td>
<td>-.19**</td>
<td>-.21**</td>
</tr>
<tr>
<td><strong>Total Social and Emotional competencies</strong></td>
<td>.24**</td>
<td>.54**</td>
<td>.45**</td>
<td>-.21**</td>
<td>1</td>
<td>.75**</td>
<td>.79**</td>
</tr>
<tr>
<td><strong>Self-awareness</strong></td>
<td>.07</td>
<td>.31**</td>
<td>.21**</td>
<td>-.15**</td>
<td>.77**</td>
<td>1</td>
<td>.47**</td>
</tr>
<tr>
<td><strong>Self-management and motivation</strong></td>
<td>.18**</td>
<td>.40**</td>
<td>.33**</td>
<td>-.05</td>
<td>.82**</td>
<td>.58**</td>
<td>1</td>
</tr>
<tr>
<td><strong>Social awareness and prosociality</strong></td>
<td>.34**</td>
<td>.50**</td>
<td>.48**</td>
<td>-.25**</td>
<td>.89**</td>
<td>.57**</td>
<td>.67**</td>
</tr>
<tr>
<td><strong>Responsible decision making</strong></td>
<td>.17**</td>
<td>.42**</td>
<td>.34**</td>
<td>-.20**</td>
<td>.75**</td>
<td>.43**</td>
<td>.48**</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01, Females above the diagonal, males below the diagonal
Figure 1. Confirmatory Factor Analysis of the Basic Empathy Scale in Poland

Note: Primary/Middle Education coefficients