Research article

Psychometric properties of the Adverse Childhood Experiences Abuse Short Form (ACE-ASF) among Romanian high school students

Franziska Meinck, Alina Paula Cosma, Christopher Mikton, Adriana Baban

**A B S T R A C T**

Child abuse is a major public health problem. In order to establish the prevalence of abuse exposure among children, measures need to be age-appropriate, sensitive, reliable and valid. This study aimed to investigate the psychometric properties of the Adverse Childhood Experiences Questionnaire Abuse Short Form (ACE-ASF). The ACE-ASF is an 8-item, retrospective self-report questionnaire measuring lifetime physical, emotional and sexual abuse. Data from a nationally representative sample of 15-year-old, school-going adolescents (n = 1733, 55.5% female) from the Romanian Health Behavior in School-Based Children Study 2014 (HBSC) were analyzed. The factorial structure of the ACE-ASF was tested with Exploratory Factor Analysis (EFA) and confirmed using Confirmatory Factor Analysis (CFA). Measurement invariance was examined across sex, and internal reliability and concurrent criterion validity were established. Violence exposure was high: 39.7% physical, 32.2% emotional and 13.1% sexual abuse. EFA established a two-factor structure: physical/emotional abuse and sexual abuse. CFA confirmed this model fitted the data well ($\chi^2$(df) = 60.526(19); RMSEA = 0.036; CFI/TLI = 0.990/0.986). Metric invariance was supported across sexes. Internal consistency was good (0.83) for the sexual abuse scale and poor (0.57) for the physical/emotional abuse scale. Concurrent criterion validity confirmed hypothesized relationships between childhood abuse and health-related quality of life, life satisfaction, self-perceived health, bullying victimization and perpetration, externalizing and internalizing behaviors, and multiple health complaints. Results support the ACE-ASF as a valid measure of physical, emotional and sexual abuse in school-aged adolescents. However, the ACE-ASF combines spanking with other types of physical abuse when this should be assessed separately instead. Future research is needed to replicate findings in different youth populations and across age groups.
1. Background

Violence against children is a major public health problem affecting an estimated one billion children annually (Hillis, Mercy, Amobi, & Kress, 2016). In particular, the experience of physical, emotional and sexual abuse and neglect in childhood is associated with poor long-term physical, mental and reproductive health outcomes (Kessler et al., 2010; Leeb, Lewis, & Zolotor, 2011; Norman et al., 2012; Paolucci, Genuis, & Violo, 2001), reduced academic performance, social and cognitive functioning, and changes in brain development (Case & Ardington, 2006; Kim & Cicchetti, 2009; Teicher, Samson, Anderson, & Ohashi, 2016). Violence exposure in childhood is also associated with a number of high-risk behaviors such as smoking, alcohol and drug use, and sexual risk behaviors, which in turn increase risk for cancer and other non-communicable diseases and sexually transmitted infections (Danese et al., 2009; Jewkes, Dunkle, Nduna, Jama, & Puren, 2010; Miller, Chen, & Parker, 2011). Furthermore, violence exposure in childhood increases the risk for perpetration and re-victimization throughout the life span and the intergenerational transmission of violence (Capaldi, Knoble, Shortt, & Kim, 2012; Coid et al., 2001; Maker, Kemmelmeier, & Peterson, 2001; Widom, Czaja, & Dutton, 2008; Widom et al., 2015).

To estimate the number of children exposed to violence in a society and monitor trends in exposure, data on incidence and prevalence is needed (Sethi et al., 2013). To date, there are only a limited number of countries that collect such data on a regular basis, including the Netherlands, the United States and Canada (Euser et al., 2013; Public Health Agency of Canada, 2010; Sedlak, 2001). Different countries adopt a variety of methods to collect data on child maltreatment. An increasing number of studies ask participants directly about the abuse or neglect they may have experienced in childhood.

These self-report studies follow four main formats: 1) large-scale cross-sectional studies with an exclusive focus on childhood victimization, its perpetrators, some socio-demographic data and a small number of outcomes (e.g., Violence Against Children Surveys [Centers for Disease Control and Prevention, 2015] or Optimus Studies [UBS Optimus Foundation, 2015]); 2) large-scale repeated cross-sectional studies which almost exclusively focus on the prevalence of child abuse, such as the National Society for the Prevention of Cruelty to Children (NSPCC) studies in the UK (May-Chahal & Cawson, 2005; Radford et al., 2011); 3) longitudinal studies that follow a cohort of children and re-interview the same children at different ages (LONGSCAN Consortium, 1990); and 4) multi-country health surveys that include a small number of items on child abuse and multiple other questionnaire components on other health problems repeated at regular intervals with children of the same age group as previous years (e.g., Health Behavior in School-aged Children or Demographic and Health Survey (Currie et al., 2014; ICF International, 1984)). These multi-country, multi-component studies have several advantages: large, nationally representative samples; regular occurrence so trends can be detected; lower-cost; and the involvement of multiple different interest groups due to the variety of topics covered. However, data on the validity and reliability of the child abuse measures used in each of these approaches is limited (Meinck and Steinert, 2015; Meinck et al., 2016a).

One of the most commonly used, non-proprietary instruments to measure child abuse exposure in surveys is the Adverse Childhood Experiences Questionnaire (ACE) (CDC, 1997; Felitti et al., 1998). The ACE Questionnaire measures the domains of physical, emotional and sexual abuse; neglect; domestic, community, and peer violence; and various dimensions of household dysfunction, such as growing up in a household where substance abuse, mental illness, or parental separation or divorce occurred (Dube, Williamson, Thompson, Felitti, & Anda, 2004; Felitti et al., 1998). It has mostly been used in the United States and other High Income Countries (HICs) but has also been used in Low and Middle Income Countries (LMICs) (Baban et al., 2013; Raleva, Peshevska, & Sethi, 2013). Thus far, psychometric testing has only been carried out on a shorter, 11-item version of the ACE which established a three-factor structure: physical/emotional abuse, household dysfunction and sexual victimization. It has good construct validity and adequate internal consistency (Ford et al., 2014).

There is also a short, 10-item screening version of the ACE that covers 10 dimensions with single-item questions. Internal consistency of the 10-item measure and construct validity were good, showing high correlations with mental and physical health measures and childhood trauma inventories (Wingenfeld et al., 2011). Given the wide cultural range in which the ACE questionnaire has been used, however, knowledge about its psychometric properties is still limited.

In response to the use of ACE in various cultural contexts, WHO developed the ACE-International Questionnaire (ACE-IQ, 43 items) to measure ACEs around the world, including low- and middle-income countries, and to be integrated into broader health surveys (World Health Organization, 2012). Although the ACE-IQ has been used in several countries, e.g. Kenya (Goodman, Martinez, Keiser, Gitari, & Seidel, 2017), Brazil (Soares et al., 2016), Saudi Arabia (Almuneef, Qayad, Aleissa, & Albuhairan, 2014), Iraq (Al-Shawi & Lafta, 2015) and Vietnam (Tran, Dunne, Vo, & Luu, 2015), data on its psychometric properties are currently also very limited.

In addition to the measures mentioned above, a short form of the ACE questionnaire (ACE-ASF) with eight items focusing solely on abuse was developed in 2012 by the World Health Organization (WHO) for use in studies with adolescents and adults. This measure does not include items on household dysfunction. No information is currently available on the psychometric properties of this instrument. This current study therefore has two aims: 1) to assess the psychometric properties of the ACE-ASF, and 2) to validate the ACE-ASF within a representative sample of Romanian youth.

2. Methods

The analysis is based on data collected as part of the 2014 “Health Behavior in School-Aged Children (HBSC)” survey in Romania. The HBSC is a WHO collaborative, cross-sectional study of adolescent health behaviors, well-being and social environment. Cross-sectional surveys of 11-, 13- and 15-year-old adolescents are carried out every four years in line with a standardized research protocol.
that specifies questionnaire content, translation guidelines and sampling methods. During each survey round, the 44 countries included in the survey (as of 2013–14) must collect data from a nationally representative sample of the three age categories discussed above (Currie et al., 2014).

Procedure: Prior to starting data collection, a pilot study of the whole national HBSC questionnaire was conducted in autumn 2014. For the age category of interest for the current study (15-year-olds), the participants included in the pilot were selected from two high schools in a big urban area. After filling out the overall HBSC 2014 questionnaire, the participants were encouraged to comment on the order of questions, the questionnaire length, and the topics covered, and to discuss whether they considered them age appropriate and age relevant. No particular concerns were raised around the abuse items.

For the main study, a nationally representative sample of students aged 11, 13 and 15 (N = 4925) enrolled in the Romanian pre-university school system was drawn from 150 schools in spring 2015 using systematic cluster sampling (schools). The sample was stratified by administrative district (county) and type of school (high school or elementary school). Directors of education from each county were asked to approve administration of the survey in their county. In the next step, principals of the randomly selected schools were approached and invited to take part in the survey (91.4% response rate). Then, active consent was sought from students and passive consent was obtained from parents. Trained research assistants and field operators administered the questionnaires following a standardized procedure in classrooms during school hours. On average, the students needed 40 min to complete the whole survey. Students were assured of the confidentiality and anonymity of their answers. As the study was completed anonymously, no referrals to child protective services could be made; however, students were provided with contact details for child helplines.

Ethics statement: The study received ethical approval from the Ethical Commission of Babes Bolyai University, Cluj Napoca, Romania.

2.1. Participants

For the purpose of this analysis, only the representative sample of the Romanian 15-year-old students was included (N = 1733, mean age 15.1, SD = 0.4, 55.5% girls and 44.6% boys), as this was the only age group asked the abuse questions.

2.2. Measures

Physical, emotional and sexual child abuse victimization were measured using a short version of the Adverse Childhood Experiences (ACE) Questionnaire which solely measures abuse, the ACE-ASF. This instrument was designed to measure lifetime exposure to emotional, physical, and sexual abuse using eight items. Emotional abuse was measured with two items that asked whether a participant had been shouted at or threatened. Physical abuse was measured with two items that covered spanking, hitting, slapping, kicking and hitting with an object. Sexual abuse was measured using four items inquiring about unwanted sexual fondling; having to touch someone sexually against one’s will; attempted unwanted oral, anal or vaginal intercourse; and unwanted oral, anal or vaginal intercourse. Respondents were provided with the following response options: (1) “I refuse to answer;” (2) “Never,” (3) “Once,” (4) “A few times,” and (5) “Many times.” To be able to carry out multi-group confirmatory factor analyses, the response options “a few times” and “many times” were combined.

Bullying victimization was assessed using the question, “How often have you been bullied at school in the past couple of months?” Bullying perpetration was assessed using the question, “How often have you bullied others at school in the past couple of months?” Questions were preceded by a preamble developed for the Olweus Bully/Victim Questionnaire (BVQ) (Muthén and Muthén, 2015; Olweus, 1996), which defines the concept of bullying and has been used in multiple other studies with similar age groups (Elgar, Craig, Boyce, Morgan, & Vella-Zarb, 2009). Respondents were asked to indicate the frequency with which they were bullied or had bullied others at school in the past couple of months with the following response categories: (1) “I haven’t;” (2) “once or twice;” (3) “two or three times in the last week;” (4) “about once a week,” and (5) “several times a week.” A dichotomous variable was created: 0 – I have not bullied others/1 have not been bullied, 1 – I have bullied others/1 have been bullied.

Perceived life satisfaction was measured using the Cantril Ladder Score across Europe and North America (Levin & Currie, 2014). Participants were shown a ladder with the following pre-amble: “Here is a picture of a ladder—the top of the ladder, 10, is the best possible life for you and the bottom is the worst possible life. In general, where on the ladder do you feel you stand at the moment?” A score of 6 or more was defined as a positive level of life satisfaction.

Self-rated health was measured using a single item used in multiple cycles of the HBSC (Galán et al., 2013): “Would you say that your health is: excellent, good, fair or bad?” Good health was defined as excellent or good health.

Psychosomatic health complaints were assessed using the HBSC symptom checklist previously used in surveys with similar populations (Gariepy, McKinnon, Sentenac, & Elgar, 2016). This instrument measures eight psychosomatic symptoms: headache, stomach ache, backache, feeling low, irritability, feeling nervous, difficulty sleeping, and feeling dizzy. Participants were asked to indicate the frequency with which they experienced the aforementioned symptoms in the past six months with option responses ranging from (1) “about every day” to (5) “rarely or never”. Suffering from more than one symptom at least once a week was coded as having multiple health complaints.

2 The ACE-ASF differs from other questionnaires in that it includes spanking as part of a range of physically abusive behaviors whereas other ACE measures either do not measure spanking or include it as a separate item.
Health-related quality of life was measured using the KIDSCREEN-10, a ten-item scale that aims to establish adolescent activity levels, stressful feelings, participation in social activities, interaction between adolescents and caregivers, adolescents’ peer relationships and self-perception of cognitive capacity and school performance measured on a scale from (1) “not at all/never” to (9) “always/extremely,” depending on the phrasing of the question. These were summed and Rasch person parameters were assigned to each possible sum score with higher values indicating better health related quality of life. These were then transformed to a mean of 50 and a standard deviation of 10. The measure has been previously used in similar populations showing good internal consistency (Ravens-Sieberer et al., 2010).

Internalizing and externalizing behaviors were measured using the widely validated Strength and Difficulties Questionnaire (SDQ) (Goodman, Meltzer, & Bailey, 1998), which shows good psychometric properties in similar populations (Mieloo et al., 2012). It consists of 25 items that measure emotional symptoms (5 items), conduct problems (5 items), hyperactivity/inattention (5 items), peer relationship problems (5 items) and prosocial behavior (5 items). These dimensions can be further grouped into internalizing problems (emotional and peer problems scale) and externalizing problems (conduct and hyperactivity problems). For the purposes of this study, we used the dimensions for both internalizing and externalizing problems.

Socio-demographics were also measured. A checklist on which participants indicated people living in their main or only home measured relationships within the home. This included mother, father, step-mother (or father’s partner), step-father (or mother’s partner), siblings, grandparents, and adults other than parents such as foster parents or care homes. Respondents were recoded as living with both parents, step-family, single parent or other. Family affluence was measured using the Family Affluence Scale III (FAS-III) (Hartley, Levin, & Currie, 2016), which shows good psychometric properties (Torsheim et al., 2016). This is a six-item assessment of common material assets or activities. Responses are summed to form an FAS score. The scale can be used as an indicator of absolute and relative socioeconomic position in society. For the purposes of this study, the relative FAS scores identified the bottom 20% (low affluence), middle 60% (medium affluence) and the highest 20% (high affluence).

2.3. Analysis

Analysis followed seven steps. First, missing data on the child abuse items were examined using SPSS 22. A small number of children (2.7%) selected the “refuse to answer” option for all abuse questions. Thus, comparisons using chi-square and independent sample t-test were conducted between those who refused to answer on all abuse questions and those who answered all or some of the abuse questions. Children who refused to answer all abuse questions were then removed from the analysis. All those who refused to answer on single items were coded as missing responses, resulting in a final sample of (n = 1668). Second, socio-demographic factors were explored and, where data were available, compared to the Romanian population (INS, 2012; Statistical Office of the European Communities, 2014). Third, Exploratory Factor Analysis (EFA) using MPlus 7 tested to examine the factor structure of the ACE-SF using geomin rotation. Fourth, to validate the results of the EFA, Confirmatory Factor Analysis (CFA) tested whether the preferred factor structure of the ACE-SF identified by the EFA fits the model well. Model fit was assessed via multiple goodness-of-fit measures. Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) above 0.95, and Root Mean Square Error of Approximation (RMSEA) indicate good model fit (Byrne, 2012). \( \chi^2 \) was also noted. As data were categorical, the weighted least square means and variance adjusted (WLSMV) estimator was used. Pairwise deletion was used to deal with missing values in all EFA and CFA analyses. Fifth, measurement invariance (configural, metric and scalar) was tested for boys and girls using Multigroup CFA (MGCFAs). WLSMV estimation using Theta parameterization was used to estimate all models; therefore model fit statistics describe the fit of the item factor model to the polychoric correlation matrix (used with categorical observed variables) among items for each group (Muthén & Muthén, 2015). The DIFFTEST procedure was used for comparisons between nested models. Configural invariance requires that the number of factors and factor loadings be the same across both groups and no equality constraints are imposed on the parameters. Thus, group differences cannot be determined. Metric invariance requires that factor loadings are constrained equally across groups, but intercepts are allowed to differ. Scores of items that meet the criteria for metric invariance can be meaningfully compared across both sexes. Scalar invariance requires that factor loadings and intercepts are constrained equal across groups. It implies that cross-sex differences in the thresholds of the observed items are due to differences in thresholds of the underlying constructs (Byrne, 2012). Sixth, internal consistency (Cronbach’s \( \alpha \)) was used to assess reliability. Finally, concurrent criterion validity was assessed through associations between the ACE-ASF physical/emotional abuse and sexual abuse constructs and correlates identified in earlier reviews and meta-analyses of quantitative research (e.g. life satisfaction, mental health, bullying victimization) (Maniglio, 2009; Norman et al., 2012).

3. Results

3.1. Respondents vs non-respondents

Of 1733 recruited adolescents, 19 failed to complete the questionnaire. Of the completed questionnaires, 2.7% (n = 46) of adolescents refused to respond to all of the abuse questions. Using \( \chi^2 \) tests, no differences could be observed between refusers and non-refusers with regards to bullying perpetration and victimization, suffering from multiple health complaints, life satisfaction and poor overall health. A significant difference could only be established for sex, with boys more likely not to respond than girls (4.3% vs 1.4%, \( \chi^2 = 14.26, p < 0.001 \)). Using independent t-test, no significant differences could be found between responders and non-responders with regards to SDQ internalizing problems scores. A significant difference was found for SDQ externalizing problems scores with higher scores among non-respondents. All participants who did not complete their questionnaires and those who refused
to respond to all of the abuse questions were removed from the analyses, resulting in a sample of n = 1668 adolescents.

3.2. Socio-demographic characteristics

Participants were 56.2% female and had a mean age of 15.1 years. 39.7% had experienced emotional abuse, 32.2% physical abuse and 13.1% sexual abuse. Girls were more likely to report physical and emotional abuse and multiple health complaints. Boys were more likely to report bullying victimization and perpetration, life satisfaction, good health and health-related quality of life. Participants were more likely to be female than the general 15-year-old population of Romania at the time of the survey and more likely to live in single-parent households compared to national Romanian statistics. All sociodemographic characteristics can be found in Table 1.

3.3. Exploratory Factor Analysis (EFA)

Examination of the model fit statistics suggested that a two-factor model using geomin rotation fit the data best (CFI 0.994, TLI 0.987, RMSEA 0.034, SRMR 0.034). The two factors identified were physical/emotional abuse and sexual abuse. Factor loadings of indicator variables were high for both, varying between 0.686–0.902 for physical/emotional abuse and 0.858–0.961 for sexual abuse (Table 3).

3.4. Confirmatory Factor Analysis (CFA)

CFA was run on the two-factor model identified by the EFA, with all physical and emotional abuse items constrained to load onto a physical/emotional abuse factor and all sexual abuse items constrained to load onto a sexual abuse factor. Results of the full model
Table 2
Summary of item phrasing, response options and frequencies for child abuse indicators that were included in the Exploratory and Confirmatory Factor Analysis for the ACE-ASF (n = 1668).

<table>
<thead>
<tr>
<th>Abuse constructs</th>
<th>Indicator</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Never</td>
</tr>
<tr>
<td>Emotional Abuse</td>
<td>‘Did a parent, guardian or other household member yell, scream or swear at you, insult or humiliate you?’</td>
<td>877 (52.6%)</td>
</tr>
<tr>
<td></td>
<td>‘Did a parent, guardian or other household member threaten to, or actually, abandon you or throw you out of the house?’</td>
<td>1443 (86.5%)</td>
</tr>
<tr>
<td>Physical Abuse</td>
<td>‘Did a parent, guardian or other household member spank, slap, kick, punch or beat you up?’</td>
<td>1047 (62.8%)</td>
</tr>
<tr>
<td></td>
<td>‘Did a parent, guardian or other household member hit or cut you with an object, such as a stick (or cane), bottle, club, knife, whip etc.? ’</td>
<td>1536 (92.1%)</td>
</tr>
<tr>
<td>Sexual Abuse</td>
<td>‘Did someone touch or fondle you in a sexual way when you did not want them to?’</td>
<td>1447 (86.8%)</td>
</tr>
<tr>
<td></td>
<td>‘Did someone make you touch their body in a sexual way when you did not want them to?’</td>
<td>1492 (89.4%)</td>
</tr>
<tr>
<td></td>
<td>‘Did someone attempt oral, anal, or vaginal intercourse with you when you did not want them to?’</td>
<td>1505 (90.2%)</td>
</tr>
<tr>
<td></td>
<td>‘Did someone actually have oral, anal, or vaginal intercourse with you when you did not want them to?’</td>
<td>1576 (94.5%)</td>
</tr>
</tbody>
</table>
3.5. Measurement invariance

Equivalence testing of the ACE-ASF across male and female sex was then carried out. Model fit statistics for all MG CFA models are found in Table 5. First, configural invariance was established by simultaneously estimating model fit for boys and girls without parameter constraints but factor variance constrained to one and factor mean fixed to 0 in each group. Model fit statistics suggest that the factor structure of the ACE-ASF is the same for boys and girls ($\chi^2 = 59.83$). Second, a model constraining factor loadings across both groups with factor means fixed to zero and residual variances fixed to 1 was evaluated for metric invariance. The metric invariance model fitted the data equally well as the configural invariance model: DIFFTEST $\Delta \chi^2(df) = 1.74(6)$, $p < 0.942$. Thus metric invariance was assumed. Third, a model constraining factor loadings and thresholds across both groups with factor means fixed to 0 in each group. Model fit for the scalar invariance model deteriorated as compared to the metric invariance model: DIFFTEST $\Delta \chi^2(df) = 101.62(14)$, $p < 0.001$. Thus scalar invariance was rejected.

3.6. Internal consistency

Internal consistency was measured using Cronbach’s $\alpha$. $\alpha$-levels were 0.57 for the physical/emotional abuse subscale, 0.83 for the sexual abuse scale and 0.71 for the total ACE-ASF scale.

3.7. Concurrent criterion validity

Correlations testing concurrent criterion validity using factor scores of the physical/emotional abuse and sexual abuse constructs confirmed hypothesized relationships. All coefficients are standardized. Physical/emotional abuse was associated with reduced general health-related quality of life ($\beta = -0.508$, $p < 0.001$), perceived health ($\beta = -0.247$, $p < 0.001$), life satisfaction (0.57 for the physical/emotional abuse subscale, 0.83 for the sexual abuse scale and 0.71 for the total ACE-ASF scale).

Table 3
Standardized results of the ACE-ASF 2-factor model Exploratory Factor Analysis among Romanian adolescents with standardized factor loadings (n = 1668).

<table>
<thead>
<tr>
<th>Abuse constructs</th>
<th>Physical &amp; Emotional Abuse</th>
<th>Sexual Abuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item:</td>
<td>$\beta$       $SE$</td>
<td>$\beta$      $SE$</td>
</tr>
<tr>
<td>'Did a parent, guardian or other household member yell, scream or swear at you, insult or humiliate you?'</td>
<td>0.800*** 0.023</td>
<td>0.902*** 0.041</td>
</tr>
<tr>
<td>'Did a parent, guardian or other household member threaten to, or actually, abandon you or throw you out of the house?'</td>
<td>0.782*** 0.044</td>
<td>0.686*** 0.060</td>
</tr>
<tr>
<td>'Did a parent, guardian or other household member spank, slap, kick, punch or beat you up?'</td>
<td>0.902*** 0.041</td>
<td>0.873*** 0.025</td>
</tr>
<tr>
<td>'Did someone make you touch their body in a sexual way when you did not want them to?'</td>
<td>0.858*** 0.039</td>
<td>0.961*** 0.044</td>
</tr>
<tr>
<td>'Did someone attempt oral, anal, or vaginal intercourse with you when you did not want them to?'</td>
<td>0.899*** 0.037</td>
<td></td>
</tr>
<tr>
<td>'Did someone actually have oral, anal, or vaginal intercourse with you when you did not want them to?'</td>
<td>0.873*** 0.025</td>
<td></td>
</tr>
</tbody>
</table>

Model fit: RMSEA 0.034; CFI/TLI 0.994/0.987; $\chi^2(df) = 38.670 (13)***; ***p < 0.001.

CFA are presented in Table 4. Fit indices indicated that the hypothesized model fitted the data well (RMSEA = 0.036; CFI 0.990, TLI 0.986, WRMR 1.026). Standardized loadings of indicators onto latent variables were high and all >0.799. The correlation between the physical/emotional abuse latent and the sexual abuse latent was moderate ($r = 0.481$, $p < 0.001$). Additional modifications to the measurement model were not carried out due to the very adequate fit of the model.

Table 4
Standardized results of the ACE-ASF 2-factor model Confirmatory Factor Analysis among Romanian adolescents with standardized factor loadings (n = 1668).

<table>
<thead>
<tr>
<th>Abuse constructs</th>
<th>Physical/Emotional Abuse</th>
<th>Sexual Abuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item:</td>
<td>$\beta$       $SE$</td>
<td>$\beta$      $SE$</td>
</tr>
<tr>
<td>'Did a parent, guardian or other household member yell, scream or swear at you, insult or humiliate you?'</td>
<td>0.799*** 0.022</td>
<td>0.854*** 0.030</td>
</tr>
<tr>
<td>'Did a parent, guardian or other household member threaten to, or actually, abandon you or throw you out of the house?'</td>
<td>0.831*** 0.030</td>
<td>0.848*** 0.020</td>
</tr>
<tr>
<td>'Did a parent, guardian or other household member spank, slap, kick, punch or beat you up?'</td>
<td>0.854*** 0.020</td>
<td>0.892*** 0.024</td>
</tr>
<tr>
<td>'Did a parent, guardian or other household member hit or cut you with an object, such as a stick (or cane), bottle, club, knife, whip etc.'</td>
<td>0.848*** 0.035</td>
<td>0.873*** 0.025</td>
</tr>
<tr>
<td>'Did someone make you touch their body in a sexual way when you did not want them to?'</td>
<td>0.924*** 0.023</td>
<td>0.902*** 0.026</td>
</tr>
<tr>
<td>'Did someone attempt oral, anal, or vaginal intercourse with you when you did not want them to?'</td>
<td>0.924*** 0.023</td>
<td>0.924*** 0.026</td>
</tr>
<tr>
<td>'Did someone actually have oral, anal, or vaginal intercourse with you when you did not want them to?'</td>
<td>0.924*** 0.023</td>
<td>0.924*** 0.026</td>
</tr>
</tbody>
</table>

Model fit: RMSEA 0.036; CFI/TLI 0.990/0.986; $\chi^2(df) = 60.526 (19)***; ***p < 0.001.
of 15-year-old school adolescents in Romania. It is the

4. Discussion

ACE-IQ (International Questionnaire—43-items) in Nigeria (Kazeem, 2015). The German sample included a non-representative adult sample of 302 psychiatric inpatients, students and control participants from the general population. The US study fielded a representative sample of 27,545 non-institutionalized adults for their EFA sample and 57,703 non-institutionalized adults for their CFA sample. The Nigerian sample included 253 prison inmates awaiting trial. All three studies differ significantly to this present study in sample size, methods and population sample. While the US study was able to carry out EFA and CFA in separate samples, sample size limitations in this current study did not allow for this to happen. The German and Nigerian studies were able to validate the ACE questionnaire against a very similar instrument, the Childhood Trauma Questionnaire (Bernstein & Fink, 1997), showing good construct validity while the US study did not examine construct validity. This present study also did not collect additional trauma data. All of the other studies used adult retrospective self-report which may make their findings prone to recall bias (Hardt & Rutter, 2004); this present study used current self-report of ACEs among youth.

Previous studies on the psychometric properties of the ACE have utilized the ACE-Screener (ten items covering ten domains) in Germany, a shortened version of the ACE with 11 items (three domains) in the US (Ford et al., 2014; Wingenfeld et al., 2011), and the ACE-IQ (International Questionnaire—43-items) in Nigeria (Kazeem, 2015). The German sample included a non-representative adult sample of 302 psychiatric inpatients, students and control participants from the general population. The US study fielded a representative sample of 27,545 non-institutionalized adults for their EFA sample and 57,703 non-institutionalized adults for their CFA sample. The Nigerian sample included 253 prison inmates awaiting trial. All three studies differ significantly to this present study in sample size, methods and population sample. While the US study was able to carry out EFA and CFA in separate samples, sample size limitations in this current study did not allow for this to happen. The German and Nigerian studies were able to validate the ACE questionnaire against a very similar instrument, the Childhood Trauma Questionnaire (Bernstein & Fink, 1997), showing good construct validity while the US study did not examine construct validity. This present study also did not collect additional trauma data. All of the other studies used adult retrospective self-report which may make their findings prone to recall bias (Hardt & Rutter, 2004); this present study used current self-report of ACEs among youth.

4.1. Factor structure

Physical and emotional abuse loaded on one dimension while sexual abuse loaded onto a different dimension. This could be explained by the small number of items (two) which measured physical and emotional abuse respectively rather than the generally recommended three (Bowen & Guo, 2012). Furthermore, previous research has shown that physical and emotional abuse often co-occur as they are carried out in most cases by the child’s primary caregiver (Meinck et al., 2016b). In addition, risk factors for child physical and emotional abuse appear to be quite distinct from risk factors for sexual abuse victimization as sexual abuse is not usually a result of conflict and perpetrators may be more deviant (Jason, Williams, Burton, & Rochat, 1982; Smith Slep & Heyman, 2001).

In line with the previous limited research on ACEs, our findings confirm that physical/emotional abuse and sexual abuse are independent, although correlated, constructs (Ford et al., 2014). The US study using the 11-item ACE-measure found a three-factor solution with the dimensions physical/emotional abuse, sexual abuse and household dysfunction (Ford et al., 2014). In the US study, physical/emotional and sexual abuse dimensions were identified by three items, which were similar to the items in this present study. As this current study did not include household dysfunction, similarity of factor structure was only hypothesized for the measured constructs. However, the US study did not include spanking as part of the physical abuse item while this present study did. Factor loadings in the US study were very similar to the ones found in this present study, as was the correlation between the physical/emotional abuse and sexual abuse constructs. Thus, findings between these studies are very similar, although slightly different ACE questionnaires were used across diverse samples and cultural contexts.

4.2. Reliability and validity

Research on concurrent criterion validity and reliability of the ACE measure is limited. The physical/emotional abuse and sexual abuse constructs showed good criterion validity in this current study through associations with a range of previously established adverse outcomes. Surprisingly, the strength of the correlations between sexual abuse and adverse outcomes was much smaller than that of physical/emotional abuse and adverse outcomes. Similarly, a recent study investigating associations between childhood...
adversity and mental health found higher odds of clinical-level mental health problems in those who experienced childhood physical abuse compared to those who experienced sexual abuse (Kessler et al., 2010). This seems to suggest that the impact of physical and emotional abuse on mental health may be more severe than that of sexual abuse, warranting further examination. The US and German studies, which investigated construct validity of the ACE-Screener and ACE-11-item version, found correlations of similar size between ACE score, depression and health, as were found in this present study (Ford et al., 2014; Wingenfeld et al., 2011). Further, studies that validated the ACE against other childhood trauma measures found very high construct validity (Kazeem, 2015; Wingenfeld et al., 2011).

In the current study, internal consistency of the sexual abuse sub-scale and of the overall ACE-ASF scale was adequate to good, while that of the physical/emotional abuse sub-scale was poor. This is in contrast to findings from the US, where the physical/emotional abuse sub-scales, sexual abuse sub-scales and ACE scale had good internal consistency (Ford et al., 2014). This discrepancy in findings may be related to differences in item-phrasing and the inclusion of spanking in the physical abuse item. The German study found good internal consistency but small inter-item correlations between some dimensions (i.e., sexual abuse and physical abuse) and large inter-item correlations between physical and emotional abuse (Wingenfeld et al., 2011). Generally, internal consistency has been found to be high across studies in different cultural contexts using differing version of the ACE (Bruskas & Tessin, 2013; Helitzer et al., 2016; Kazeem, 2015).

4.3. Multi-group comparisons

In the measurement invariance test, the ACE-ASF demonstrated full metric invariance, suggesting that meaningful comparisons of scores of items can be made across sex. This is in line with a previous study testing the psychometric properties of an ACE measure with 11 items in the US, where full metric invariance across sex and age (Ford et al., 2014) was established. The ACE-ASF failed to demonstrate full scalar invariance; thus, cross-sex differences in the thresholds of the observed items are likely due to differences in thresholds of the underlying constructs.

4.4. Limitations

This study is subject to a number of limitations. First, although this was a nationally representative study of the Romanian population of 15-year-olds, girls and children parented by single mothers were over-represented. Second, whereas the internal consistency of the ACE-ASF scale and the sexual abuse sub-scale was adequate, the physical/emotional abuse sub-scale showed poor internal consistency. As Cronbach’s α is affected by the number of items in each sub-scale, an increase in the number of items may improve internal consistency but would also increase the item content of the measure (Tavakol & Dennick, 2011). Third, the original ACE measure includes a household dysfunction dimension which is absent from this present measure. This measure is thus a measure of child abuse—rather than adverse experiences more generally—and does not take the role of the social determinants of health into account which are well established and recognized by public health agencies (Commission on Social Determinants of Health, 2010). Fourth, the measure included an item which combined spanking with more serious physical abuse types (e.g. kick, punch or beat up). While the harmful effects of frequent spanking have been well established (Affi et al., 2017; Merrick et al., 2017), prior research distinguished between multiple instances and none or very few instances of spanking. The ACE-ASF does not give researchers the option to make this distinction and puts children who have been spanked once or twice into the same category as those who have been punched or kicked. Spanking should thus become an individual item for future research using this measure. Fifth, this study lacked diversity as it only included 15-year-old school children in Romania. Therefore, the findings of this study cannot be generalized to other age groups, ethnicities or more vulnerable populations who may have dropped out of school. However, the sample was nationally representative of 15-year-old school children, and as schooling is mandatory in Romania until age 16, the vast majority of children would have been in school. Sixth, we were unable to compare the ACE-ASF to one of the long-established measures as it was impossible to include these in the overall study. Seventh, due to small numbers of participants reporting sexual abuse, it was impossible to split the sample into two datasets, one for the EFA and one for the CFA. Future research should conduct EFA and CFA on different or split samples. Finally, the study relied on adolescent self-report of abusive events. The sensitive nature of the topic might have influenced participants’ willingness to disclose these traumatic events (Tourangeau & Yan, 2007). However, adolescent self-report has been found to be more reliable than agency records, parental report of adolescent victimization and adult retrospective self-report (Hardt & Rutter, 2004; Johnsona et al., 2002; Maier, Mohler-Kuo, Landolt, Schnyder, & Jud, 2013).

4.5. Implications for policy and practice

The results confirm that the ACE-ASF is a valid measure of physical, emotional and sexual child abuse. Even with the limitations reported in this study, the ACE-ASF is supported by more evidence for its validity than many other measures currently in use. It can be rapidly and easily completed, and is freely available. The high response rates of the ACE questionnaires in this and other surveys suggest the ACE is an acceptable measure that translates well across different cultural contexts. The ACE-ASF considers corporal punishment such as spanking as a form of child abuse, and practitioners and researchers wanting to draw a distinction between corporal punishment and abuse should contemplate using separate items for spanking versus more violent forms of abusive behavior (i.e. kicking or beating up). Due to its brevity, the ACE-ASF can be included in multi-component surveys or be used by itself in short surveys. This measure fills a significant gap for validated child abuse measures created by the current increased focus on violence against children. This is due in part to the inclusion of violence against children as an explicit target—16.2—in the Sustainable
Development Goals (SDG) and the creation of the Global Partnership to End Violence to help achieve this SDG target (United Nations General Assembly, 2015). The ACE-ASF measure gives countries the opportunity to establish the burden of child abuse within their youth population and measure changes in child abuse victimization through repeated cross-sectional measurements in a time- and cost-effective way. The resulting data can then be used to inform much-needed child protection and child abuse prevention policy and programming. The ACE-ASF is, therefore, a valuable measure for consideration in child abuse research and should be more widely used and tested.

4.6. Future research needs

The findings of this study have implications for global research on child abuse victimization. Considering the dearth of psychometric evaluations of the ACE-ASF and other child abuse screening tools, future research is urgently needed on the psychometric properties of these measures in varying cultural contexts and, in particular, in low- and middle-income countries. In addition, in-depth qualitative research on the appropriateness, cultural applicability, cognitive understanding, and willingness to respond to the questions within the measures in different cultural contexts and throughout the stages of child development is required. Further, more research is needed validating child abuse measures in adolescent samples using current child self-report and around the inclusion of spanking with more severe physical abuse types.

Author contributions

AB and AC were involved in the overall study design and management. FM and AC had responsibility for conceptualizing and writing the paper. FM led and conducted the analyses. AC, AB and CM contributed to the analyses and the interpretation of findings. All authors reviewed and approved the final version.

Acknowledgments

The study was funded by a Norway Innovation Grant [RO19.04] toward the Romanian Ministry of Health/National Institute of Public Health as part of the “Multilevel intervention to prevent non-communicable disease associated with the life style” (2014–2016). Franziska Meinck was supported by an Economic and Social Research Council (ESRC) Future Research Leader Award [ES/N017447/1], the European Research Council under the European Union’s Seventh Framework Programme [FP7/2007-2013, grant agreement 313421], the Leverhulme Trust [PLP-2014-095] and the University of Oxford’s ESRC Impact Acceleration Account [1602-KEA-189; 1311-KEA-004 & 1609-GCRF-227].

The authors wish to thank the participants for taking time to fill in the survey, Prof Peter Schmidt for his helpful and inspiring course at the Essex Summer School in Social Science Data Analysis, and Jennifer Rabebeau for proof-reading and editing.

References


